UNIVERSITY
OF MICHIGAN
MAR 15 1951
MATHEMATICS
LIBRARY

AMERICAN JOURNAL OF MATHEMATICS

FOUNDED BY THE JOHNS HOPKINS UNIVERSITY

EDITED BY

R. BRAUER
UNIVERSITY OF MICHIGAN

L. M. GRAVES
UNIVERSITY OF CHICAGO

S. EILENBERG
COLUMBIA UNIVERSITY

D. C. LEWIS, JR.
THE JOHNS HOPKINS UNIVERSITY

A. WINTNER
THE JOHNS HOPKINS UNIVERSITY

WITH THE COÖPERATION OF

L. V. AHLFORS	C. B. ALLENDOERFER	W. L. CHOW
R. ARENS	R. BAER	P. R. HALMOS
J. D. HILL	P. HARTMAN	H. SAMELSON
N. H. McCOY	J. J. STOKER	R. M. THRALL
J. L. SYNGE	O. SZÁSZ	H. WEYL

PUBLISHED UNDER THE JOINT AUSPICES OF

THE JOHNS HOPKINS UNIVERSITY
AND
THE AMERICAN MATHEMATICAL SOCIETY

Volume LXXII, Number 4
OCTOBER, 1950

THE JOHNS HOPKINS PRESS
BALTIMORE 18, MARYLAND
U. S. A.

CONTENTS

	PAGE
Free sums of groups and their generalizations II. By REINHOLD BAER,	625
Free sums of groups and their generalizations III. By REINHOLD BAER,	647
Generalized free sums of cyclical groups. By HANNA NEUMANN,	671
Generalized Euclidean space in terms of a quasi inner product. By LEONARD M. BLUMENTHAL,	686
On the spectrum of Hilbert's matrix. By WILHELM MAGNUS,	699
The expansion theorem for sigma-monogenic functions. By LIPMAN BERS,	705
On the spectrum of a boundary value problem with two singular end-points. By KENNETH G. WOLFSON,	713
On cube-coverings of n-space. By L. G. PECK,	720
On the local uniqueness of geodesics. By PHILIP HARTMAN,	723
On the Hölder restrictions in the theory of partial differential equations. By AUREL WINTNER,	731
On the zeta-functions of algebraic number fields II. By RICHARD BRAUER,	739
On completeness in the sense of Archimedes. By L. W. COHEN and CASPER GOFFMAN,	747
On completeness and category in uniform space. By L. W. COHEN and CASPER GOFFMAN,	752
On the fundamental equations of differential geometry. By PHILIP HARTMAN and AUREL WINTNER,	757
Lamé coordinates in Hilbert space. By PHILIP HARTMAN and AUREL WINTNER,	775
An application of the theory of basic series to theorems of Bernstein-Widder type. By PHILIP DAVIS,	787
The mean convergence of orthogonal series. By G. MILTON WING,	792
On the structure of some subgroups of the modular group. By EMIL GROSSWALD,	809
On the solution of sum-equations. By I. M. SHEFFER,	835
The gaps in the essential spectra of wave equations. By PHILIP HARTMAN and CALVIN R. PUTNAM,	849
On non-uniformly quasi-analytic functions. By I. I. HIRSCHMAN, JR.,	863

The AMERICAN JOURNAL OF MATHEMATICS will appear four times yearly.

The subscription price of the JOURNAL for the current volume is $7.50 (foreign postage 50 cents); single numbers $2.00.

A few complete sets of the JOURNAL remain on sale.

Papers intended for publication in the JOURNAL may be sent to any of the Editors.

Editorial communications should be sent to Professor AUREL WINTNER at The Johns Hopkins University.

Subscriptions to the JOURNAL and all business communications should be sent to THE JOHNS HOPKINS PRESS, BALTIMORE 18, MARYLAND, U. S. A.

Entered as second-class matter at the Baltimore, Maryland, Postoffice, acceptance for mailing at special rate of postage provided for in Section 1103, Act of October 3, 1917, Authorized on July 3, 1918.

PRINTED IN THE UNITED STATES OF AMERICA
BY J. H. FURST COMPANY, BALTIMORE, MARYLAND

FREE SUMS OF GROUPS AND THEIR GENERALIZATIONS II.*

By Reinhold Baer.

Recently we have shown that the appropriate framework for a discussion of the existence of free sums with amalgamations is provided by the concept of add and that one has to determine those adds which are amalgams of their subgroups and satisfy the strong associative law. Since the latter law is stated in terms of vectors over the add, the problem arises of giving an intrinsic characterization; and this problem we are going to solve in the present investigation [Section **4**]. An application of this result is then given in Section **5**.

For our purposes [and also for later applications] it proves convenient to analyze the concept of amalgam of subgroups and to break it down into formal laws concerning the addition of elements. These formal laws are the "postulates" of Section **1**; and in Sections **2** and **3** we show that they really characterize amalgams of subgroups. In order to obtain these results we have to solve the problem [of interest in itself] to characterize those subsets of an add which are contained in a subgroup of the add [Section **2**].

1. The postulates. An *add* is a set A of elements subject to a composition which we call addition and which meets only the following requirement:

The sum of the elements a and b in A is either undefined or else it is a uniquely determined element $a+b$ in A.

We note the possibility that the sum $a+b$ is defined while the sum $b+a$ is undefined.

The adds we shall be interested in will always be of a more special type. We shall therefore enumerate a number of formal requirements. These postulates will not be independent. But we shall require at various stages of our investigation the validity of various sets of these postulates; and thus this abundance will prove a convenience.

Postulates of Inversion.

I. *There exists an element* 0 *in* A *such that* $x+0=0+x=x$ *for every element* x *in* A.

It is clear that 0 is uniquely determined by Postulate I.

* Received June 6, 1949.

II. *To every x in A there exists y in A such that $x + y = 0$.*

III. *$x + y = 0$ and $x + z = 0$ imply $y = z$.*

If I, II and III are satisfied by A, then there exists to every element x in A one and only one element $-x$ in A such that $x + (-x) = 0$. This symbol $-x$ will be used whenever we are assured of the validity of Postulates I, II, III; and then we shall write $x - y$ instead of $x + (-y)$ etc.

IV. *$x + y = 0$ implies $y + x = 0$.*

Postulates I to IV imply $-(-x) = x$. Adds which meet these four requirements have been termed elsewhere self-reflexive [Baer (1), Section 4].

V. *$x + (y + z) = 0$ if, and only if, $(x + y) + z = 0$.*

This is in reality a rather elaborate statement. It asserts that the existence of the sums $y + z$ and $x + (y + z)$ and the validity of the equation $x + (y + z) = 0$ are equivalent to the existence of the sums $x + y$ and $(x + y) + z$ together with $(x + y) + z = 0$. The following proposition will be useful.

(0) *If Postulates* I *and* IV *are satisfied, then the following properties of the add A are equivalent.*

(i) *Postulate* V.

(ii) *$x + (y + z) = 0$ implies $(x + y) + z = 0$.*

(iii) *$(x + y) + z = 0$ implies $x + (y + z) = 0$.*

Proof. It clearly suffices to show that (ii) implies (iii). Hence assume the validity of (ii) and suppose that $(x + y) + z = 0$. Then we deduce from IV and (ii) successively the validity of the following equations: $z + (x + y) = 0$, $(z + x) + y = 0$, $y + (z + x) = 0$, $(y + z) + x = 0$, $x + (y + z) = 0$, so that (iii) is true.

POSTULATES OF ASSOCIATIVITY.

VI. *If the sums $x + y$, $y + z$, $(x + y) + z$, $x + (y + z)$ exist in A, then $(x + y) + z = x + (y + z)$.*

VII. *If the sums $x + y$ and $y + z$ exist in A, then the existence of $(x + y) + z$ is necessary and sufficient for the existence of $x + (y + z)$.*

It is worth noting that the antecedents of the postulates V, VI, VII require the existence of 2, 4, 3 sums respectively. We note the validity of the following important rule.

(1) *Assume validity of* I, VI, VII *and existence of* $r+s$. *Then* $q+r=0$ *implies* $q+(r+s)=s$ *and* $s+t=0$ *implies* $(r+s)+t=r$.

This is an immediate consequence of $s=(q+r)+s$ and $r=r+(s+t)$.

It follows from (1) that adds satisfying Postulates I to VII are self-reflexive in the strict sense [Baer (1), Section **4**].

(2) *If Postulates* I, V, VI, VII *are satisfied by* A, *if* $x+y$ *exists in* A, *and if* $x+x'=y+y'=0$, *then* $y'+x'$ *exists in* A *and satisfies* $(x+y)+(y'+x')=0$.

Proof. We infer from (1) that $x=(x+y)+y'$. Consequently $0=[(x+y)+y']+x'$; and our contention is a consequence of V.

(3) *Postulates* III, IV *are consequences of* I, II, VI.

Proof. Assume $x+y=0$. There exists by II an element z such that $y+z=0$. It follows from I that $x=x+(y+z)$, $z=(x+y)+z$; and consequently we may infer from VI that $x=x+(y+z)=(x+y)+z=z$. Thus IV is a consequence of I, II, VI.

If $x+y=x+z=0$, then we infer from IV that $z+x=0$; and it follows as before that $y=(z+x)+y=z+(x+y)=z$. Hence III is a consequence of I, IV, VI and therefore of I, II, VI.

(4) *Postulate* V *is a consequence of* I, II, VI, VII.

Proof. Assume that $x+(y+z)=0$. There exists by II an element w in A such that $z+w=0$. Hence it follows from (1) that $y=(y+z)+w$; and we have by I that $w=[x+(y+z)]+w$. Now it follows from VI, VII that

$$w=[x+(y+z)]+w=x+[(y+z)+w]=x+y.$$

From IV and $z+w=0$ we infer now that $0=w+z=(x+y)+z$. Thus we have shown that $x+(y+z)=0$ implies $(x+y)+z=0$.

From this fact and IV we deduce successively: if $(r+s)+t=0$, then $t+(r+s)=0$, $(t+r)+s=0$, $s+(t+r)=0$, $(s+t)+r=0$, $r+(s+t)=0$; and this completes the proof of the validity of V.

We notice that all the postulates I to VII are valid whenever we are assured of the validity of I, II, VI, VII. We shall make much use of this fact.

POSTULATES OF EXISTENCE.

VIII. *The existence of the sums* $x+y$, $y+z$, $x+z$ *implies the existence of* $(x+y)+z$.

In conjunction with VII this implies also the existence of $x+(y+z)$.

(5) *If Postulates* I, V, VI, VII, VIII *are satisfied by* A, *if the sums* $x+x$, $x+y$, $y+y$ *exist in* A, *and if there exist elements* x', y' *such that* $x+x'=x'+x=y+y'=y'+y=0$, *then the sums* $x+y'$ *and* $y+x$ *exist in* A.

Proof. We infer from (2) the existence of $y'+y'$ and from (1) the validity of $x=(x+y)+y'$. From VIII and the existence of the sums $(x+y)+y'$, $y'+y'$, $(x+y)+y'$ we infer the existence of

$$[(x+y)+y']+y'=x+y'.$$

From $x+x'=y'+y=0$ and the existence of $x+y'$, just established, we infer by (2) the existence of $y+x'$. But then it follows from our hypotheses and the result of the first paragraph that $y+x$ exists too, as we desired to show.

IX. *The sum* $x+y$ *exists in* A *if, and only if,* $y+x$ *exists in* A.

X. *The sum* $x+x$ *exists for every* x *in* A.

(6) *Postulate* IX *is a consequence of Postulates* I, II, VI, VII, VIII *and* X.

This is a fairly immediate consequence of (3), (4), (5).

We shall show in the next section that these last three postulates assure the existence of sufficiently many subgroups of the add A; and this is the reason for terming these postulates the Postulates of Existence. We note that these last three postulates have considerably more depth than the preceding ones. It is for this reason that we shall try to avoid using them as much as possible.

2. The subgroups of an add. Throughout this section we shall consider adds satisfying Postulate I, so that we may speak always of the zero element in A. Every subset of the add A may be considered as a subadd of A. If a subadd of A contains 0 and is a group with respect to the addition defined in A, then we shall term it *a subgroup of* A. Our main problem in this section is the determination of those subsets of A which are contained in subgroups of A. With this in mind we state

DEFINITION 1. *The subset* S *of the add* A *is a C-set, if*

(C. 1) *the sum* $x+y$ *exists in* A, *whenever* x *and* y *are in* S; *and*

(C. 2) *to every* s *in* S *there exists an element* s' *in* S *such that* $s+s'=0$.

PROPOSITION 1. *If Postulates* I, VI, VII, VIII *are satisfied by the add* A, *then every* C*-set is contained in a subgroup of* A.

Proof. Suppose that S is a C-set in the add A. Denote by $S(0)$ the subset of A which consists of 0 and the elements in S. Clearly $S(0)$ is a C-set too. We define now inductively subsets $S(i)$ of A by the following rule:

(0) $S(i+1) = S(i) + S(i)$ consists of all the sums $x+y$ for x, y in $S(i)$ which exist in A.

We prove now inductively:

(1) $S \leq S(0) \leq S(1) \leq \cdots \leq S(i) \leq S(i+1) \leq \cdots$.

(2) Every $S(i)$ is a C-set.

The assertion (1) is an immediate consequence of the fact that $T \leq T+T$, provided 0 is in T. To prove (2) we note first that $S(0)$ is a C-set. Hence we may make the inductive hypothesis that $S(i)$ is a C-set. If x and y are elements in $S(i)$, then their sum $x+y$ exists in A [by (C. 1)] and belongs to $S(i+1)$. We infer from (C. 2) the existence of elements x', y' in $S(i)$ such that $x+x' = y+y' = 0$. It follows from (C. 1) that the sum $y'+x'$ of the elements y' and x' in $S(i)$ exists in A and belongs to $S(i+1)$. We infer from § **1**, (1) that $x = (x+y)+y'$. Hence $0 = [(x+y)+y']+x'$. Now we infer from the existence of $y'+x'$ in A and from Postulates VI, VII that

$$0 = [(x+y)+y']+x' = (x+y)+(y'+x');$$

and this shows the validity of (C. 2) in $S(i+1)$. If x, y, z are elements in $S(i)$, then we infer from (C. 1) the existence of the sums $x+y$, $y+z$, $x+z$ in A; and now we may infer from Postulate VIII the existence of $(x+y)+z$ in A. If a, b, c, d are elements in $S(i)$, then we infer from the preceding remark and from (C. 1) the existence of the sums $(a+b)+c$, $c+d$, $(a+b)+d$ in A; and consequently we may deduce from Postulate VIII the existence of $[(a+b)+c]+d$. But $c+d$ exists in A [by (C. 1)]; and now we infer from Postulate VII the existence of the sum $(a+b)+(c+d)$ in A. This shows the validity of (C. 1) in $S(i+1)$. Hence $S(i+1)$ is a C-set too; and this completes the inductive proof of (2).

Denote by G the set-theoretical join of the sets $S(i)$. It follows from (1) that $S \leq G$. If x and y are elements in G, then there exists by (1) an integer m such that x and y are both in $S(m)$. It follows from (2) that $S(m)$ is a C-set so that the sum $x+y$ exists in A. Hence $x+y$ is in

$S(m+1)$ by (0): and consequently $x+y$ is in G. Thus we see that the sum of elements in G is a uniquely determined element in G; and it follows from Postulate VI that addition in G is associative. From $S(0) \leq G$ we infer that 0 is in G. If finally x belongs to G, then x belongs to some $S(k)$; and it follows from (2) that $S(k)$ contains an element x' such that $x+x'=0$. Hence there exists to every x in G an element x' in G such that $x+x'=0$; and now it is easily verified that G is a subgroup of the add A which contains S, as we intended to prove.

PROPOSITION 2. *If the Postulates* I, V, VI, VII, VIII *are satisfied by the add A, then the following conditions are necessary and sufficient for the subset S of A to be part of a subgroup of the add A.*

(a) *If x and y are in S, then the sum $x+y$ exists in A.*

(b) *To every x in S there exists an element x' in A such that $x+x'=x'+x=0$.*

Proof. The necessity of the conditions (a) and (b) is obvious. Assume now that the subset S of A satisfies the conditions (a) and (b). Then we may select to every element s in S some element s' in A such that $s+s'=s'+s=0$; and we may form the set T of all the elements s and s' for s in S. It is clear that Condition (C. 2) is satisfied by T. If x and y are elements in S, then we infer from Condition (a) the existence of the sums $x+x$, $x+y$, $y+y$, $y+x$ in A. Hence we may deduce from §1, (2) and §1, (5) the existence of the sums $x'+x'$, $y'+x'$, $y'+y'$, $y'+x'$ and $x+y'$, $y'+x$. This proves the validity of (C. 1) in T. Hence T is a C-set; and it follows from Proposition 1 that T is contained in a subgroup of the add A. But S is part of T; and so S itself is part of a subgroup of the add A.

COROLLARY 1. *Assume the validity of the postulates* I, II, VI, VII, VIII *in the add A. Then*

(a) *the subset S of A is part of a subgroup of A if, and only if, $x+y$ exists in A whenever x and y are in S.*

(b) *Every subgroup of A is part of a maximal subgroup of A.*

(c) *If M is a maximal subgroup of A, if w is an element in A, but not in M such that $w+w$ exists in A, then there exists an element v in M such that at least one of the sums $w+v$ and $v+w$ does not exist in A.*

Proof. We infer from §1, (3) and §1, (4) the validity of all the postulates I to VIII in A. Now (a) is an immediate consequence of Proposition 2.

The contention (b) is verified by the customary set-theoretical arguments [Maximum Principle]. Suppose now that G is a subgroup of the add A, and that w is an element in A such that the sums $w+w$, $w+g$, $g+w$ exist for every g in G. If S is the set composed of w and the elements in G, then the sum $x+y$ exists in A whenever x and y are in S. It follows from (a) that S is part of a subgroup of A. If G were in particular a maximal subgroup of A, then this would imply that w is an element in G; and this proves (c).

DEFINITION 2. *The add A is the amalgam of its subgroups [in the strict sense], if*

(A. 1) *every element in A is contained in a subgroup of A;*

(A. 2) *cross cuts of subgroups of A are subgroups of A;*

(A. 3) *the existence of the sums $x+y$, $y+z$, $z+x$ in A implies the existence of a subgroup of A which contains x, y, z.*

Amalgams in the weak sense will be defined in the next section. [See also Baer (1), §6.]

COROLLARY 2. *If the Postulates* I, II, VI, VII, VIII *are satisfied by the add A, and if the subadd $E=E(A)$ consists of all the elements e in A such that the sum $e+e$ exists in A, then $E(A)$ has the following properties.*

(a) *If x is in E, and if x' is an element in A such that $x+x'=0$, then x' belongs to E.*

(b) *If x and y are in E, and if the sum $x+y$ exists in A, then $x+y$ belongs to E.*

(c) *Every subgroup of A is part of E.*

(d) *E is the amalgam of its subgroups [in the strict sense].*

Proof. We note first that [by §1, (3) and §1, (4)] all Postulates I to VIII are satisfied by A. If $x+x$ exists in A, and if $x+x'=0$, then it follows from §1, (2) that $x'+x'$ exists in A; and this proves (a). If $x+x$, $x+y$, $y+y$ exist in A, then we infer from §1, (5) the existence of the sum $y+x$ in A. It follows from Corollary 1, (a) that there exists a subgroup of A which contains x and y; and this implies in particular the existence of the sum $(x+y)+(x+y)$. This proves (b). (c) is trivial. It is a consequence of Corollary 1, (a) that the element x in A belongs to a subgroup of A if, and only if, $x+x$ exists in A. Hence (A. 1) is satisfied by E.

That (A. 2) is satisfied by E is a consequence of the fact that the equation $a + x = 0$ possesses for every a in A one and only one solution x in A. Consider finally elements x, y, z in E such that the sums $x + y, y + z, z + x$ exist in A. Then it follows from §1, (5) that the sums $y + x, z + y, x + z$ exist in A; and the sums $x + x, y + y, z + z$ exist in A, since x, y, z are in E. Hence it follows from Corollary 1, (a) that there exists a subgroup S of A which contains x, y, z; and that $S \leq E$, is obvious [and has been pointed out before]. Thus E is the amalgam of its subgroups, as we claimed.

THEOREM. *The add A is the amalgam of its subgroups [in the strict sense] if, and only if, Postulates* I, II, VI, VII, VIII, X *are satisfied by A.*

Proof. It is a consequence of Corollary 2, (d) that A is an amalgam, if Postulates I, II, VI, VII, VIII, X are satisfied by A, since X implies $A = E(A)$. Assume conversely that A is the amalgam of its subgroups [in the strict sense]. It is a consequence of (A. 2) that all subgroups of A have the same null-element; and now we deduce Postulate I from (A. 1). Postulates II and X are likewise consequences of (A. 1). If the sum $x + y$ exists in A, then we infer from (A. 3) and the existence of the sums $0 + x$, $x + y$, $y + 0$ that $0, x, y$ are contained in some subgroup of A. Thus we have verified:

(A. 3*) *The sum $x + y$ exists in A if, and only if, x and y are contained in some subgroup of A.*

If now the sums $x + y, y + z, z + x$ exist in A, then it follows from (A. 3*) that the sum $x + z$ exists likewise. It follows from (A. 3) that x, y, z are contained in some subgroup of A; and thus $(x + y) + z$ exists in A. This shows the validity of VIII.

If the sums $x + y, y + z, (x + y) + z$ exist in A, then it follows from (A. 3*) that x, y belong to one subgroup and that y, z belong to some subgroup of A. Thus we are assured of the existence of the sums $y + (x + y)$, $(x + y) + z, z + y$; and we infer from (A. 3) the existence of a subgroup S of A which contains $y, x + y, z$. But it is clear that x, y, z belong to the subgroup S of A. Now one sees readily the validity of the following property.

(A. 3**) *The following properties of the elements x, y, z in the amalgam A are equivalent.*

(i) *There exists a subgroup of A which contains x, y, z.*

(ii) *The sums $x + y, y + z, z + x$ exist in A.*

(iii) *The sums* $x+y$, $y+z$, $(x+y)+z$ *exist in* A.

(iv) *The sums* $x+y$, $y+z$, $x+(y+z)$ *exist in* A.

From (A. 3**) one deduces readily the validity of Postulates VI and VII. This completes the proof.

Remark. It is a consequence of §1, (3), (4), (6) that all Postulates I to X are satisfied by amalgams.

3. Amalgams satisfying the strong associative law. The characterization of amalgams which we obtained in the preceding section can be simplified considerably, if we assume that the add under consideration satisfies the strong associative law. In order to state this law in the form best suited for our purposes some concepts have to be recalled.

If $a_1, \cdots, a_n$, $0 < n$, are elements in the add A, then we term the ordered n-tuplet $(a_1, \cdots, a_n)$ a vector of length n over A [and sometimes it will simplify notations to make use of the vector of length 0 too which is the empty set]. If $(a_1, \cdots, a_n)$ is a vector a over A, and if the sum $a_i + a_{i+1}$ exists in A, then we term the vector $(a_1, \cdots, a_{i-1}, a_i + a_{i+1}, a_{i+2}, \cdots, a_n)$ of length $n-1$ a *contraction* of the vector a of length n. If the vector b arises from the vector a by a sequence of contractions, then b is termed a *reduction* of a [note that every vector is its own reduction too]. Finally we say that the vectors u and v are similar $[u \sim v]$, if there exist vectors $w(0), \cdots, w(i), \cdots, w(2k)$ such that $u = u(0)$, $w(2k) = v$ and such that $w(2i)$ and $w(2i+2)$ are reductions of $w(2i+1)$. [For a discussion of the significance of these concepts see Baer (1); but note that we have slightly changed the terminology.]

SIMILARITY CRITERION. *If Postulates* I *to* VII *are satisfied by the add* A, *then similarity of the vectors* u *and* v *is a necessary and sufficient condition for the existence of a vector* w *such that* u *and* v *are reductions of* w.

For a proof see Baer (1), §4, Corollary 1, where for Postulates I to VII the slightly weaker hypothesis has been substituted that A be strictly self-reflexive.

THE STRONG ASSOCIATIVE LAW *holds in the add* A, *if*

(S. 1) $(a) \sim (b)$ *implies* $a = b$ *and*

(S. 2) *similarity of the irreducible vector* v *to a vector of length* 1 *implies that* v *too be of length* 1.

Here we term a vector irreducible, if it does not possess any reductions. It has been shown [Baer (1), §3, Theorem 1] that this formulation of the strong associative law is equivalent with the one originally proposed [see Baer (1) for a discussion of the importance of this law.]

LEMMA 1. *The strong associative law implies Postulates* VI *and* VII.

Proof. If the sums $x+y$, $y+z$ and $(x+y)+z$ exist in A, then we have the vector similarities

$$((x+y)+z) \sim (x+y, z) \sim (x, y, z) \sim (x, y+z).$$

Now it follows from (S. 2) that the last vector is reducible. Hence there exists the sum $x+(y+z)$ in A. Consequently the vectors $((x+y)+z)$ and $(x+(y+z))$ are similar over A; and we infer $(x+y)+z=x+(y+z)$ from (S. 1). If the sums $x+y$, $y+z$ and $x+(y+z)$ exist in A, then we show similarly the existence of $(x+y)+z$ and the equality $x+(y+z) = (x+y)+z$. This shows the validity of VI and VII.

DEFINITION. *The add A is the amalgam of its subgroups [in the weak sense], if*

(B. 1) *every element in A belongs to some subgroup of A;*

(B. 2) *cross cuts of subgroups of A are subgroups of A;*

(B. 3) *the existence of the sum $x+y$ in A implies the existence of a subgroup of A which contains both x and y.*

LEMMA 2. *The add A is an amalgam in the strict sense if, and only if,*

(a) *A is an amalgam in the weak sense and*

(b) *to any given subgroups X, Y, Z of A there exists a subgroup S of A which contains the cross cuts $X \cap Y$, $Y \cap Z$, $Z \cap X$.*

Proof. Assume that A is an amalgam in the strict sense. Then it is easily seen that A is an amalgam in the weak sense too [we have verified this in §2, when we deduced (A. 3*)]. Consider now subgroups X, Y, Z of the add A and element r, s which belong to the join of $X \cap Y$, $Y \cap Z$ and $Z \cap X$. Then r and s belong to one of the subgroups X, Y, Z; and thus their sum $r+s$ exists in A. Since in A all Postulates I to X are satisfied [§2, Theorem], we may apply §2, Corollary 1, (a); and this proves the existence of a subgroup S of A which contains $X \cap Y$, $Y \cap Z$, $Z \cap X$. Thus the conditions (a) and (b) are necessary. If conversely the conditions (a)

and (b) are satisfied by A, then A certainly satisfies (A. 1), (A. 2). If x, y, z are elements such that the sums $x+y$, $y+z$, $z+x$ exist in A, then we infer from (a) and (B. 3) the existence of subgroups X, Y, Z of A such that x, y belongs to X, y, z to Y and z, x to Z. But then y is in $X \cap Y$, z in $Y \cap Z$ and x in $Z \cap X$. It follows from (b) that there exists a subgroup S of A which contains these cross cuts and which therefore contains the elements x, y, z. Thus (A. 3) is satisfied by A too; and this completes the proof.

On the basis of this lemma it is easy to construct amalgams in the weak sense which are not amalgams in the strict sense.

THEOREM. *If the strong associative law is satisfied by the add A, then the following properties of A are equivalent.*

(i) *A is an amalgam in the strict sense.*

(ii) *A is an amalgam in the weak sense.*

(iii) *Postulates* I, II, IX, X *are satisfied by A.*

(iv) *Postulates* I, II, VIII, X *are satisfied by A.*

Proof. We have mentioned before [Lemma 2] that (i) implies (ii): and that amalgams in the weak sense satisfy Postulates I, II, IX, X, is easily deduced from the definition.

Assume now the validity of Postulates I, II, IX, X. We infer from Lemma I [and the strong associative law] the validity of Postulates VI, VII so that we may deduce from §1, (3), (4) the validity of Postulates III, IV, V. Thus we may use the symbol $-x$ in the usual fashion. We prove first:

IX*. *The existence of $x+y$ implies the existence of $x-y$.*

If $x+y$ exists in A, then there exist the sums $y+x$, $-y-y$, $x+x$.

Consequently the following vector similarities are true:

$$(x+y, -y-y, y+x) \sim (x, y, -y, -y, y, x) \sim (x, x) \sim (x+x).$$

This implies by (S. 2) the reducibility of the first of these vectors; and consequently at least one of the sums $(x+y)+(-y-y)$ and $(-y-y)+(y+x)$ exists in A. From §1, (1) we infer $x=(x+y)-y$ and $x=-y+(y+x)$; and now we deduce from VI, VII the existence of $(x+y)+(-y-y)=[(x+y)-y]=y=x-y$ or of $(-y-y)+(y+x)=-y+[-y+(y+x)]=-y+x$. But it

follows from IX that $x - y$ exists if, and only if, $-y + x$ exists; and this completes the proof of IX*.

Now we are ready to verify VIII. Assume the existence of the sums $x + y$, $y + z$, $x + z$ in A. This implies [by IX*] the existence of $x - y$, $y - z$, $x - z$. Consequently we have the vector similarities:

$$(x - y, y - z) \sim (x, -y, y, -z) \sim (x, -z) \sim (x - z);$$

and we infer from the strong associative law that

$$(x - y) + (y - z) = x - z.$$

Thus it follows from IX that the sums $(y - z) + (x - y)$ and $-z + x$ exist in A; and this implies the validity of the following vector similarities:

$$((y - z) + (x - y)) \sim (y, -z, x, -y) \sim (y, -z + x, -y).$$

Hence it follows from (S. 2) that the last of these vectors is reducible; and consequently at least one of the sums $y + (-z + x)$ and $(-z + x) - y$ exists in A; and it follows from IX and IX* that both these sums exist in A. Using the equation $-(-t) = t$ one deduces now successively from VII, IX, IX* the existence of the sums

$$(-z + x) + y, -z + (x + y), (x + y) - z, (x + y) + z.$$

Thus VIII is satisfied by the add A; and we have shown that (iv) is a consequence of (iii).

If Postulates I, II, VIII, X are satisfied by A, then we see as before that Postulates VI, VII are satisfied too; and it follows from §2, Theorem that A is an amalgam in the strict sense. This completes the proof.

4. Intrinsic characterization of the strong associative law. It is our objective in this section to characterize by inner properties the adds which satisfy the strong associative law. We precede this characterization by the proot of a couple of useful lemmas.

LEMMA 1. *Assume that Postulates* I *and* V *are satisfied by the add* A.
(a) *If* (0) *is a reduction of the vector* $(v_1, \cdots, v_n)$, *and if* $2 < n$, *then there exists an integer* i *with the following properties:*

$0 < i < n - 1$, *the sum* $v_i + v_{i+1}$ *exists in* A *and* (0) *is a reduction of the contracted vector* $(v_1, \cdots, v_{i-1}, v_i + v_{i+1}, v_{i+2}, \cdots, v_n)$.

(b) (0) *is a reduction of the vector* $(v_1, \cdots, v_n)$ *with* $1 < n$ *if, and only if, there exists a reduction* (a) *of* $(v_1, \cdots, v_{n-1})$ *such that* $a + v_n = 0$.

Proof. If (0) is a reduction of the vector (a, b, c) of length 3, then either $(a+b)+c=0$ or $a+(b+c)=0$. It is a consequence of Postulate V that these two equations are equivalent conditions on the elements a, b, c. Consequently the sum $a+b$ exists in A; and (0) is a reduction of $(a+b, c)$. Thus Proposition (a) has been verified for $n=3$. Assume now that $3<n$ and that Proposition (a) has been verified for $n-1$. If (0) is a reduction of the vector $(v_1, \cdots, v_n)$, then there exists certainly an integer k with the following properties.

$0<k<n$, the sum v_k+v_{k+1} exists in A, and (0) is a reduction of the contracted vector $w=(\cdots, v_{k-1}, v_k+v_{k+1}, v_{k+2}, \cdots)$.

If k happens to be less than $n-1$, then k is the desired integer; and if $k=n-1$, then (0) is a reduction of the vector w whose length $n-1$ is at least 3. Thus we may apply the inductive hypothesis on

$$w=(v_1, \cdots, v_{n-2}, v_{n-1}+v_n).$$

Hence there exists an integer i with the following properties:

$0<i<n-2$, the sum v_i+v_{i+1} exists in A, and (0) is a reduction of the [twice contracted] vector

$$t=(v_1, \cdots, v_{i-1}, v_i+v_{i+1}, v_{i+2}, \cdots, v_{n-2}, v_{n-1}+v_n).$$

But then (0) is a reduction of the vector

$$s=(v_1, \cdots, v_{i-1}, v_i+v_{i+1}, v_{i+2}, \cdots, v_n),$$

of which t is a [simple] contraction. Thus we have found a suitable integer i in either case. This completes the induction and the proof of (a). (b) is readily deduced from (a) by complete induction.

Remark. By interchanging right and left in Lemma 1 another pair of correct assertions is obtained.

LEMMA 2. *If Postulates* I, IV *and* V *are satisfied by the add A, and* i_r^s (0) *is a reduction of the vector* $(v_1, \cdots, v_n)$ *with* $1<n$, *then* (0) *is also a reduction of the vector* $(v_i, \cdots, v_n, v_1, \cdots, v_{i-1})$.

Proof. Clearly it suffices to prove that (0) is a reduction of $(v_n, v_1, \cdots, v_{n-1})$ whenever (0) is a reduction of $(v_1, \cdots, v_n)$. But if (0) is a reduction of $(v_1, \cdots, v_n)$, then there exists, by Lemma 1, (b), a reduction (w) of $(v_1, \cdots, v_{n-1})$ such that $w+v_n=0$. It follows from Postulate IV that $0=v_n+w$. Hence (0) is a reduction of (v_n, w) and this vector is clearly a reduction of $(v_n, v_1, \cdots, v_{n-1})$ so that (0) is also a reduction of this latter vector, as we desired to prove.

THEOREM. *If Postulates* I *and* II *are satisfied by the add* A, *then the following properties of* A *are equivalent.*

(i) *The strong associative law.*

(ii) *Postulate* VI *and Property* (S. 2).

(iii) *Postulates* VI, VII *together with the following property*:

(S) *If the sums* $x_1 - x_2, \cdots, x_{n-1} - x_n$, $x_n - x_1$ *exist in* A, $3 < n$, *then the sum* $x_{i-1} - x_{i+1}$ *exists in* A *for some* i *such that* $1 < i < n$.

(iv) *Postulates* III, IV, V *together with the following property*:

(S. 0) *If* (0) *is a reduction of the vector* $(v_1, \cdots, v_n)$ *with* $1 < n$, *and if the difference* $v_{n-1} + v_n$ *exists in* A, *then* (0) *is also a reduction of the contracted vector* $(v_1, \cdots, v_{n-2}, v_{n-1} + v_n)$.

We note that the Postulates III, IV, V are consequences of the Postulates I, II, VI, VII [see Section 1] and that it is therefore possible to use in Condition (iii. S) the symbol $-x$.

Proof. It is an immediate consequence of §3, Lemma 1 that (i) implies (ii). Assume next the validity of (ii). Suppose that the sums $a + b$, $b + c$ exist in A, and that at least one of the sums $(a + b) + c$ and $a + (b + c)$ exists in A. If, for instance, $a + (b + c)$ exists in A, then we have the vector similarities:

$$(a + (b + c)) \sim (a, b, c) \sim (a + b, c).$$

Since the first of these vectors has length 1 and the last one length 2, we infer from (S. 2) reducibility of the last of these vectors. Hence the sum $(a + b) + c$ exists in A; and this proves the validity of VII. Thus, I, II, VI, VII are satisfied by A; and it follows as usual that III, IV, V hold too. Hence we may use the symbol $-t$. Suppose now that the sums $x_1 - x_2, \cdots, x_{n-1} - x_n$, $x_n - x_1$ exist in A. Then we infer from §1, (2) the existence of $-(x_n - x_1) = -(-x_1) - x_n = x_1 - x_n$; and consequently we have the following vector similarities:

$$(x_1 - x_2, \cdots, x_{n-1} - x_n) \sim (x_1, -x_2, x_2, -x_3, \cdots, x_{n-1}, -x_n)$$
$$\sim (x_1, -x_n) \sim (x_1 - x_n).$$

The first of these vectors has length $n - 1$ and the last one length 1. Thus we infer from $2 < n$ and (S. 2) the reducibility of the first of these vectors. Consequently there exists an integer i such that $0 < i < n - 1$ and such that the sum $(x_i - x_{i+1}) + (x_{i+1} - x_{i+2})$ exists in A. This implies the validity of the following vector similarities:

$$((x_i - x_{i+1}) + (x_{i+1} - x_{i+2})) \sim (x_i, -x_{i+1}, x_{i+1}, -x_{i+2})$$
$$\sim (x_i, -x_{i+2}).$$

But then (S. 2) implies the reducibility of the last of these vectors; and this shows the existence of the sum $x_i - x_{i+2}$. Thus we have verified (S); and this shows that (iii) is a consequence of (ii).

Assume now the validity of (iii). Then we see as usual the validity of all the Posulates I to VII. Property (S. 0) is certainly satisfied by vectors of length 2; and it is equivalent with V for vectors of length 3; and thus we may make the inductive hypothesis that (S. 0) is satisfied by all vectors whose length is less than n and that $3 < n$. Suppose now that (0) is a reduction of the vector $(v_1, \cdots, v_n)$ and that the sum $v_{n-1} + v_n$ exists in A. We distinguish two possibilities.

Case 1. There exists an integer i with the following properties:

$0 < i < n-2$, the sum $v_i + v_{i+1}$ exists in A, and (0) is a reduction of the contracted vector $(v_1, \cdots, v_{i-1}, v_1 + v_{i+1}, v_{i+2}, \cdots, v_n)$.

Since this contracted vector has length $n-1$, we may apply the inductive hypothesis. Since $i+1 < n-1$, this implies that (0) is a reduction of the twice contracted vector

$$(v_1, \cdots, v_{i-1}, v_i + v_{i+1}, v_{i+2}, \cdots, v_{n-2}, v_{n-1} + v_n).$$

But then it is clear that (0) is likewise a reduction of the [once contracted] vector $(v_1, \cdots, v_{n-2}, v_{n-1} + v_n)$, as we claimed.

Case 2. If $0 < i < n-2$, and if the sum $v_i + v_{i-1}$ exists in A, then (0) is not a reduction of the contracted vector

$$(v_1, \cdots, v_{i-1}, v_i + v_{i+1}, v_{i+2}, \cdots, v_n).$$

[Incidentally this can only happen for $n > 4$.]

In this case we deduce from Lemma 1, (a) inductively the existence of elements $w_{n-1}, \cdots, w_1$ with the following properties:

(a) the vector (w_i) is a reduction of the vector $(v_i, v_{i+1}, \cdots, v_{n-1})$;

(b) $w_{i-1} = v_{i-1} + w_i$;

(c) (0) is a reduction of the vector $(v_1, \cdots, v_{i-1}, w_i, v_n)$.

We note first two immediate consequences:

(a′) $w_{n-1} = v_{n-1}$;

(a″) $(v_1, \cdots, v_{i-1}, w_i, v_n)$ is a reduction of the vector $(v_1, \cdots, v_n)$.

From §1, (1) and from (b) we deduce $v_{i-1} = (v_{i-1} + w_i) - w_i = w_{i-1} - w_i$; and thus we have by §1, (2):

(b′) the sum $w_i - w_{i-1}$ exists in A for $1 < i < n$.

From (c) we deduce that (0) is a reduction of (w_1, v_n). Hence $0 = w_1 + v_n$ or

(c′) $w_1 = -v_n$.

Finally we deduce from the existence of $v_{n-1} + v_n$ the existence of $-(v_{n-1} + v_n) = -v_n - v_{n-1} = w_1 - w_{n-1}$ by §1, (2) and by (c′), (a′). Thus we have shown that the following sums exist in A:

$$w_1 - w_{n-1}, w_{n-1} - w_{n-2}, \cdots, w_2 - w_1.$$

It follows now from Condition (S) that either the sum $w_1 - w_{n-2}$ or one of the sums $w_{i+1} - w_{i-1}$ with $1 < i < n-1$ exists in A. [This is trivial, if $n = 4$.]

We show next that the second possibility cannot arise. For otherwise some sum $w_{i+1} - w_{i-1}$ with $1 < i < n-1$ would exist in A; and we would infer from §1, (2) that $w_{i-1} - w_{i+1}$ exists in A. It follows from (b) that

$$w_{i-1} - w_{i+1} = (v_{i-1} + w_i) - w_{i+1}.$$

But it follows from (b) [and §1, (1)] that $v_i = w_i - w_{i+1}$; and now we deduce from Postulates VI, VII that

$$\begin{aligned} w_{i-1} &= (w_{i-1} - w_{i+1}) + w_{i+1} = [(v_{i-1} + w_i) - w_{i+1}] + w_{i+1} \\ &= [v_{i-1} + (w_i - w_{i+1})] + w_{i+1} = (v_{i-1} + v_i) + w_{i+1}. \end{aligned}$$

Now we deduce from (c) that (0) is a reduction of the vector

$$(v_1, \cdots, v_{i-2}, w_{i-1}, v_n) = (v_1, \cdots, v_{i-2}, (v_{i-1} + v_i) + w_{i+1}, v_n);$$

and this vector is a contraction of the vector

$$(v_1, \cdots, v_{i-2}, v_{i-1} + v_i, w_{i+1}, v_n).$$

But it follows from (a) that the last vector is a reduction of

$$(v_1, \cdots, v_{i-2}, v_{i-1} + v_i, v_{i+1}, \cdots, v_{n-1}, v_n),$$

so that (0) has been shown to be a reduction of this vector contrary to the hypothesis of Case 2.

Thus we have shown the existence of the sum $w_1 - w_{n-2}$. Consequently we may deduce from (c′) the existence of the sum

$$-(w_1 - w_{n-2}) = w_{n-2} - w_1 = w_{n-2} + v_n.$$

Since the sum $v_{n-1} + v_n$ exists in A, we may now deduce from (b), (a′) and the Postulates VI, VII that

$$w_{n-2} + v_n = (v_{n-2} + w_{n-1}) + v_n = (v_{n-2} + v_{n-1}) + v_n$$
$$= v_{n-2} + (v_{n-1} + v_n).$$

It follows from (c) that (0) is a reduction of the vector

$$(v_1, \cdots, v_{n-3}, w_{n-2}, v_n)$$

whose length is $n - 1$. Since the sum $w_{n-2} + v_n$ exists in A, it follows from the inductive hypothesis that (0) is likewise a reduction of the vector

$$(v_1, \cdots, v_{n-3}, w_{n-2} + v_n) = (v_1, \cdots, v_{n-3}, v_{n-2} + (v_{n-1} + v_n)).$$

But this vector is a contraction of $(v_1, \cdots, v_{n-3}, v_{n-2}, v_{n-1} + v_n)$ so that (0) is a reduction of this last vector, as we desired to show. This completes the inductive proof of (S. 0); and thus we have shown that (iv) is a consequence of (iii).

Assume now the validity of (iv). Then we prove first:

(iv. 1) If (0) is a reduction of the vector $v = (v_1, \cdots, v_n)$, then (0) is a reduction of every contraction of v.

If the sum $v_i + v_{i+1}$ exists in A, $0 < i < n$, then we deduce first from Lemma 2 that (0) is a reduction of the vector

$$v' = (v_{i+2}, \cdots, v_n, v_1, \cdots, v_i, v_{i+1});$$

and we deduce from Condition (S. 0) that (0) is a reduction of the contraction

$$v'' = (v_{i+2}, \cdots, v_n, v_1, \cdots, v_{i-1}, v_i + v_{i+1});$$

and consequently it follows from Lemma 2 again that (0) is a reduction of the contraction $(v_1, \cdots, v_{i-1}, v_i + v_{i+1}, x_{i+2}, \cdots, v_n)$ of v, as we desired to show.

(iv. 2) Suppose that $v \sim w$. Then (0) is a reduction of v if, and only if, (0) is a reduction of w.

To prove this we note first the following immediate consequence of (iv. 1): If the vector r is a contraction of the vector s, then (0) is a reduction of r if, and only if, (0) is a reduction of s; and from this fact the proposition (iv. 2) follows by complete induction.

Suppose now that $(a_1, \cdots, a_n)$ is an irreducible vector over A of positive length n; and that a is an element in A such that

$$(a_1, \cdots, a_n) \sim (a).$$

Since the Postulates I to V are satisfied by A, we may make use of the operation $-t$. Consequently we find that

$$(0) \sim (a, -a) \sim (a_1, \cdots, a_n, -a) = v.$$

It follows from (iv. 2) that (0) is a reduction of the vector v; and now it follows from Lemma 1, (b) and the irreducibility of the vector $(a_1, \cdots, a_n)$ that $n = 1$ and $a_1 - a = 0$. But this implies [because of Postulates I to V] that $a_1 = a$. Thus we have shown the validity of the strong associative law; and this completes the proof.

COROLLARY 1. *The add A is an amalgam satisfying the strong associative law if, and only if, A satisfies Postulates* I, II, VI, VII *and*

(S*) *the existence of the sums $x_1 + x_2, \cdots, x_{n-1} + x_n, x_n + x_1$ with $3 < n$ implies the existence of a sum $x_{i-1} + x_{i+1}$ with $1 < i < n$.*

Proof. If A is an amalgam satisfying the strong associative law, then A satisfies Postulates I, II, VI, VII [§3, Lemma 1]; and the existence of $a + b$ implies clearly the existence of $a - b$. Thus the existence of the sums $x_1 + x_2, \cdots, x_{n-1} + x_n, x_n + x_1$ implies the existence of the differences $x_1 - x_2, \cdots, x_{n-1} - x_n, x_n - x_1$; and it follows from the preceding Theorem that at least one of the differences $x_{i-1} - x_{i+1}$ exists in A. But this implies the existence of the sum $x_{i-1} + x_{i+1}$ in the amalgam A. Thus we have shown the necessity of our conditions.

Assume conversely the validity of Postulates I, II, VI, VII and (S*). If $a + b$ exists in A, then we are assured of the existence of the following sums: $a + b$, $b - b\ [= 0]$, $-b - a\ [= -(a + b)]$, $-a + a$; and it follows from (S*) that at least one of the sums $a - b$ and $b - a$ exists in A. But the existence of $b - a$ implies the existence of $-(b - a) = a - b$; and thus we have shown that the existence of $a + b$ implies the existence of $a - b$. This Property IX* implies together with (S*) the property (S) of the preceding Theorem, so that the strong associative law holds in A. One deduces furthermore from IX* as usual the Postulates IX and X; and now it follows from §3, Theorem that A is an amalgam in the strict sense satisfying the strong associative law.

COROLLARY 2. *The amalgam A [in the weak sense] satisfies the strong associative law, if*

(a) *to any three distinct maximal subgroups X, Y, Z of A there exists a subgroup S of A such that $X \cap Y$, $Y \cap Z$, $Z \cap X$ are all part of S; and*

(b) *to any $n > 3$ maximal subgroups $X_1, \cdots, X_{n-1}, X_n = X_0$ of A there exists a subgroup T of A and an integer i with $0 < i < n-1$ such that $X_{i-1} \cap X_i$ and $X_{i+1} \cap X_{i+2}$ are both part of T.*

Proof. It is a consequence of (a) and §3, Lemma 2 that A is an amalgam in the strict sense; and we deduce from (b) that Corollary 1, (S*) is satisfied by A. Thus the validity of the strong associative law may be inferred from Corollary 1.

Remark 1. It is an open question whether Condition (b) is necessary, though certain parts of (b) may be shown to be necessary. It is easy, however, to construct examples which show that it does not suffice to require the validity of (b) for bounded n only.

Remark 2. Most of the known criteria for the validity of the strong associative law are easily seen to be special cases of Corollary 2; see in particular Baer (1) §8, Theorem 1 and Corollary 1.

5. Generalization of H. Neumann's reduction theorem. We suppose throughout this section that the add A satisfies Postulates I to V. Then A possesses a uniquely determined null element, the operation $-x$ is well defined, and cross cuts of subgroups are subgroups of A.

If M is a maximal subgroup of A, then we form the subgroup $C(M)$ of M which is generated by all the subgroups $M \cap X$, where X ranges over all the maximal subgroups, not M, of A. The join $D(A)$ of all these subgroups $C(M)$ is a subadd of A with a number of important properties:

$$C(M) = M \cap D(A) \text{ for every maximal subgroup } M \text{ of } A.$$

Consequently $D(A)$ contains $-x$, whenever it contains x; and it contains $x + y$, whenever $x + y$ exists in A and x and y belong to $D(A)$. Hence $D(A)$ is an amalgam [in the weak sense]; and it is an amalgam in the strict sense whenever A has this property [see also Baer (1), §8, Lemma 3]. It can also be shown that the strong associative law is satisfied by the amalgam A if, and only if, it is satisfied by $D(A)$ [Baer (1), §8, Theorem 2; Neumann (1), 5.0. Theorem p. 599]; and it is our objective to prove an analogous property for the iterates of $D(A)$ which includes the result just mentioned as a special case.

We define the iterates of the operation D in the following obvious fashion:

$$A = D_0(A),$$

$D_{\nu+1}(A) = D[D_\nu(A)]$,

$D_\sigma(A($ = intersection of all the $D_\nu(A)$ with $\nu < \sigma$ in case σ is a limit ordinal.

One proves easily by complete [transfinite] induction that the $D_\nu(A)$ form a descending chain of subadds of A with the closure property:

$D_\nu(A)$ contains $-x$ whenever it contains x; and it contains $x+y$ whenever it contains x and y and the sum $x+y$ exists in A.

The $D_\nu(A)$ with $0 < \nu$ are all amalgams of their subgroups; and if A is an amalgam in the strict sense, then all the $D_\nu(A)$ are amalgams in the strict sense.

Finally we note the existence of an ordinal τ such that $D_\tau(A) = D_{\tau+1}(A) = D[D_\tau(A)]$. It is of interest to note that there exist [easily constructed] amalgams A in the strict sense where this ordinal τ cannot be finite.

THEOREM. *If A is an amalgam in the strict sense, then the following properties of A are equivalent.*

(i) *The strong associative law is satisfied by A.*

(ii) *The strong associative law is satisfied by every $D_\nu(A)$.*

(iii) *The strong associative law is satisfied by at least one $D_\nu(A)$.*

Proof. We note first that every $D_\nu(A)$ is an amalgam in the strict sense. Assume now the validity of the strong associative law in A; and consider elements $a_1, \cdots, a_n$ in $D_\rho(A)$ with the property: the sums $a_1 + a_2, \cdots, a_{n-1} + a_n$, $a_n + a_1$ exist in $D_\rho(A)$ and $3 < n$. Since the elements a_i and their sums are in A, and since the strong associative law holds in A, we may deduce from §4, Corollary 1, (S*) the existence of one of the sums $a_{i-1} + a_{i+1}$ with $1 < i < n$ in A. But $D_\rho(A)$ is closed in A with respect to addition; and so this sum exists in $D_\rho(A)$. Consequently Condition (S*) is satisfied by $D_\rho(A)$; and we deduce from §4, Corollary 1 the validity of the strong associative law in $D_\rho(A)$. Thus we have shown that (ii) is a consequence of (i); and (iii) is just a special case of (ii).

Assume now the validity of (iii). Then we want to show the validity of Condition (S*) of §4, Corollary 1 in A. If this were not true, then there would exist a minimal integer $m > 3$ with the following properties:

(T. m) there exist elements $w_1, \cdots, w_m$ in A such that the sums $w_1 + w_2, \cdots, w_{m-1} + w_m, w_m + w_1$ exist in A, whereas none of the sums $w_{i-1} + w_{i+1}$ with $1 < i < m$ exists in A.

We show next that these elements w_i, satisfying (T. m), belong to every

$D_\nu(A)$. This is certainly true for $\nu = 0$; and thus we may make the inductive hypothesis:

$$0 < \nu \text{ and the } w_i \text{ belong to every } D_\sigma(A) \text{ with } \sigma < \nu.$$

If ν happens to be a limit ordinal, then $D_\nu(A)$ is the intersection of all the $D_\sigma(A)$ with $\sigma < \nu$ so that the w_i belong to $D_\nu(A)$ too [because of the inductive hypothesis]. If $\nu = \rho + 1$ is not a limit-ordinal, then it follows from the inductive hypothesis that the w_i belong to $D_\rho(A)$. Since $D_\rho(A)$ is closed in A with respect to addition, it follows from (T. m) that all the sums $w_1 + w_2, \cdots, w_{m-1} + w_m, w_m + w_1$ belong to $D_\rho(A)$. Since $D_\rho(A)$ is an amalgam of its subgroups, there exist maximal subgroups S_i of $D_\rho(A)$ such that w_i, w_{i+1} are in S_i for $0 < i < m$ and such that w_m, w_1 belong to S_m. If $S_i = S_{i+1}$ for some positive $i < m - 1$, then w_i, w_{i+1}, w_{i+2} would be in the same subgroup of A so that the sum $w_i + w_{i+2}$ would exist contrary to (T. m). Hence

$$S_i \neq S_{i+1} \text{ for } 0 < i < m - 1.$$

Next we show that $S_{m-1} \neq S_m \neq S_1$. To do this we distinguish two cases.

Case 1. $m = 4$. If $S_3 = S_4$, then w_3, w_4, w_1 would be in the same subgroup of A so that the sum $w_1 + w_3$ would exist in A, contrary to (T. 4); and if $S_4 = S_1$, then w_4, w_1, w_2 would belong to the same subgroup of A so that the sum $w_2 + w_4$ would exist in A, contrary to (T. 4). Hence $S_3 \neq S_4 \neq S_1$, as we claimed.

Case 2. $4 < m$. If $S_{m-1} = S_m$, then w_{m-1}, w_m, w_1 belong to the same subgroup of A so that the sum $w_{m-1} + w_1$ exists in A; and the elements $w_1, \cdots, w_{m-1}$ satisfy (T. m − 1). If $S_m = S_1$, then w_m, w_1, w_2 belong to some subgroup of A so that the sum $w_m + w_2$ exists in A; and the elements $w_2, \cdots, w_m$ satisfy (T. m − 1). Either way we are led to a contradiction with our minimal choice of m. Hence $S_{m-1} \neq S_m \neq S_1$, as we claimed.

Since the S_i are maximal subgroups of the amalgam $D_\rho(A)$, the operation $C(S_i)$ is well defined with respect to $D_\rho(A)$. The element w_1 belongs to the cross cut $S_1 \cap S_m$ of two distinct maximal subgroups of $D_\rho(A)$. Hence w_1 is in $C(S_1)$ and therefore in $D[D_\rho(A)]$. Likewise w_i with $1 < i \leq m$ belongs to the cross cut $S_{i-1} \cap S_i$ of two distinct maximal subgroups of $D_\rho(A)$ and hence to $D[D_\rho(A)]$. Thus we have shown that every w_i belongs to $D[D_\rho(A)] = D_{\rho+1}(A) = D_\nu(A)$; and this completes our inductive proof.

We have shown now that Condition (S*) of §4, Corollary 1 is not satisfied by any $D_\nu(A)$, if it is not satisfied by A. But it follows from (iii) and §4, Corollary 1 that this Condition (S*) is satisfied by at least one

$D_\nu(A)$. Hence it is satisfied by A too; and this implies the validity of the strong associative law in A. Hence (i) is a consequence of (iii); and this completes the proof.

Remark 1. If any two distinct maximal subgroups of the amalgam A meet in the same subgroup U, then $D(A) = U$; and it follows from the preceding theorem that the strong associative law is satisfied by A. Thus one sees that Schreier's Theorem is a special case of the preceding theorem. On the other hand it does not seem likely that a criterion like §4, Corollary 2 could be deduced from our result.

Remark 2. If A is an amalgam in the strict sense, then there exists some ordinal τ such that the amalgam $D_\tau(A) = B$ is an amalgam in the strict sense which satisfies $B = D(B)$. It follows from our theorem that the strong associative law holds in A if, and only if, it holds in B. Thus it suffices to find criteria for the validity of the strong associative law in amalgams B [in the strict sense] which satisfy in addition the condition $B = D(B)$.

UNIVERSITY OF ILLINOIS,
URBANA, ILLINOIS.

BIBLIOGRAPHY.

Reinhold Baer.

(1) "Free sums of groups and their generalizations. An analysis of the associative law," *American Journal of Mathematics*, vol. 71 (1949), pp. 706-742.

B. H. Neuman and Hanna Neumann.

(1) "A remark on generalized free products," *Journal of the London Mathematical Society*, vol. 25 (1950), pp. 202-204.

Hanna Neumann.

(1) "Generalized free products with amalgamated subgroups," *American Journal of Mathematics*, vol. 70 (1948), pp. 590-625.

FREE SUMS OF GROUPS AND THEIR GENERALIZATIONS III.*

By REINHOLD BAER.

In the first communication of this series we have shown that the appropriate framework for an investigation of free sums of groups with amalgamated subgroups is provided by the concept of add [= system with an incomplete addition]. In particular we have stressed the central importance of what we called the strong associative law. This may be restated in a form appropriate for the purposes of this introduction as follows: similar irreducible vectors are equal, if one of them has length 1. The connection of such a requirement with the so-called word [or identity] problem is obvious, since it requires a trivial solution in a special case. It is our objective in the present investigation to characterize adds meeting the following still stronger requirements: (a) the class of adds where similar irreducible vectors are equal [Section **2**]; (b) the class of adds where similar irreducible vectors have equal length and where similar vectors of length 1 are equal [Section **1**]. In adds of class (a) the word problem has a trivial solution, and in adds of class (b) a simple solution of the word problem exists [Section **3**]. In Section **4** we establish finally a striking and surprising relation between the word problem and the transformation problem.

Results and concepts of our second communication are used freely in the present investigation. References to this paper are indicated by the prefix II [like II, §**4**, Lemma 1]. The postulates I to X which we enumerated in II, §**1** are used so often that we refer to them just by their roman numeral without any further reference.

1. Invariance of length of irreducible reductions. The principal objective of this section is the proof of the equivalence of three properties of adds. Before enunciating this theorem it will be convenient to discuss some of the concepts and properties occurring in this theorem.

THE ASSOCIATIVE LAW (T) consists of the Property (S. 1) of the strong associative law [or associative law (S) as we shall term it in the future] and of the following invariance property:

* Received June 6, 1949.

(L) *Similar irreducible vectors have the same length.*

It is clear that the Property (S. 2) is a special case of (L), since it requires the validity of (L), if at least one of the vectors has length 1. Thus the associative law (S) is a consequence of the associative law (T).

The property (L) is easily seen to be equivalent to the apparently weaker property

(L′) *If v' and v'' are irreducible reductions of the vector v, then v' and v'' have the same length.*

Proof of the equivalence of Properties (L) *and* (L′): if v' and v'' are reductions of v, then they are similar vectors; and thus (L′) is a consequence of (L). Suppose conversely the validity of (L′) and the similarity of the irreducible vectors u and w. Then there exist vectors $v(i)$ such that $u = v(0)$, $w = v(2n)$ and $v(2i)$, $v(2i+2)$ are reductions of $v(2i+1)$ for $0 \leq i < n$. [Note that some of these vectors may be equal.] Since u and w are irreducible vectors, and since every vector possesses an irreducible reduction, we may assume without loss in generality that the vectors $v(2j)$ are irreducible. But then it follows from (L′) that $v(2i)$ and $v(2i+2)$ for $0 \leq i < n$ have the same length; and thus u and w have the same length. Hence (L) is a consequence of (L′).

We note finally that Property (S. 1) is not a consequence of Property (L). as may be seen from the example of any system with a unique, but non-associative, addition where all irreducible vectors have length 1. That (S. 1) does not imply (S. 2) nor (L) is likewise easily seen.

THE POSTULATE XI consists of three parts:

(a) *If the sums $a+b$, $b+c$, $c+d$ exist in A, then at least one of the sums $a+(b+c)$ and $(b+c)+d$ exists in A.*

(b) *If the sums $b+c$, $c+d$ and $a+(b+c)$ exist in A, then at least one of the sums $a+b$ and $(b+c)+d$ exists in A.*

(c) *If the sums $a+b$, $b+c$ and $(b+c)+d$ exist in A, then at least one of the sums $a+(b+c)$ and $c+d$ exists in A.*

In certain instances it is possible to deduce the properties (b), (c) from (a); but whether or not this is true in general, the author does not know.

CONNECTED CONTRACTIONS. This concept is not of as great an intrinsic interest as the preceding ones; but it will prove very helpful in several proofs. We define it as follows:

If $v(1), \cdots, v(m)$ are contractions of the vector v, and if $v(i)$ and $v(i+1)$ possess, for $0 < i < m$, a common contraction, then $v(1)$ and $v(m)$ are connected contractions of v.

We shall utilize this concept for $m = 2, 3$ only. The reason for this is contained in the following useful proposition.

LEMMA. *If there exist contractions of the vector v which are not connected, then v satisfies one of the following two (exclusive) properties:*

(a) *$v = v' + (a, b, c) + v''$ where $v' + (a)$ and $(c) + v''$ are irreducible vectors and where the sums $a + b$, $b + c$ exist in A, though the existence of the sums $a + (b + c)$ and $(a + b) + c$ would imply $a + (b + c) \neq (a + b) + c$.*

(b) *$v = w' + (p, q, r, s) + w''$ where $w' + (p)$ and $(s) + w''$ are irreducible vectors and where the sums $p + q$, $q + r$, $r + s$ exist in A, though the existence of the sums $p + (q + r)$ and $(p + q) + r$ would imply $p + (q + r) \neq (p + q) + r$ and the existence of the sums $q + (r + s)$ and $(q + r) + s$ would imply $q + (r + s) \neq (q + r) + s$.*

Proof. Our arguments are based on the following obvious constructions:

(A) If $v = x + (a, b) + y + (c, d) + z$, and if the sums $a + b$ and $c + d$ exist in A, then $x + (a + b) + y + (c + d) + z$ is a common contraction of the contractions $x + (a + b) + y + (c, d) + z$ and $x + (a, b) + y + (c + d) + z$ of the vector v.

(B) If $v = x + (a, b, c) + y$, if the sums $a + b$, $b + c$, $a + (b + c)$, $(a + b) + c$ exist in A, and if $a + (b + c) = (a + b) + c = s$, then $x + (s) + y$ is a common contraction of the contractions $x + (a + b, c) + y$ and $x + (a, b + c) + y$ of the vector v.

If the vector $v = (v_1, \cdots, v_n)$ possesses contractions which are not connected, then v possesses at least two contractions. Consequently there exists an integer k such that $1 < k < n$ and integers $f(1), \cdots, f(k)$ such that $0 < f(1) < \cdots < f(j) < f(j+1) < \cdots < f(k) < n$ with the following properties:

the sum $v_i + v_{i+1}$ exists in A if, and only if, $i = f(j)$ for some j with $1 \leq j \leq k$.

If we denote now by $v(j)$ the vector which arises from v by substituting for $v_{f(j)}$, $v_{f(j)+1}$ their sum, then the vectors $v(1), \cdots, v(k)$ constitute all the contractions of v.

It follows from (A) that $v(i)$ and $v(j)$ have a common contraction whenever $f(i)+1<f(j)$. This implies in particular that $v(1)$ and $v(i)$ possess a common contraction for $2<i$ and that $v(j)$ and $v(k)$ possess a common contraction for $j<k-1$. Consequently any two contractions of v are connected if $3<k$; or if $k=3$, but $f(1)+2<f(3)$. The existence of unconnected contractions of v implies therefore that $k=2$ or 3, $f(2)=f(1)+1$ and, in case $k=3$, then $f(3)=f(2)+1=f(1)+2$. The validity of our lemma is now readily deduced from (B).

THEOREM 1. *The following properties of the add A are equivalent.*

(i) *The associative law* (T).

(ii) *The postulates* VI, VII, XI.

(iii) *If the contractions v' and v'' of the vector v are not both irreducible, then they are connected contractions of v.*

Proof. Assume first the validity of the associative law (T). This implies the associative law (S) of which VI, VII are consequences [II, §3, Lemma 1]. If the sums $a+b$, $b+c$, $c+d$ exist in A, then we have the vector similarities:

$$(a+b, c+d) \sim (a, b, c, d) \sim (a, b+c, d).$$

Hence it follows from the invariance rule (L) that the last of these vectors is reducible; and this proves the existence of at least one of the sums $a+(b+c)$ and $(b+c)+d$; i. e. the validity of XI, (a). If next the sums $a+b, b+c$, $(b+c)+d$ exist in A, then we obtain the vector similarities:

$$(a, (b+c)+d) \sim (a, b, c, d) \sim (a+b, c, d).$$

As before we deduce from (L) the reducibility of the last of these vectors and therefore the existence of at least one of the sums $(a+b)+c$ and $c+d$ in A so that the validity of XI, (c) has been established; and XI, (b) is verified similarly. Thus we have shown that (ii) is a consequence of (i).

Assume now the validity of (ii). Then we prove (iii) in the following slightly stricter form:

(iii*) *If the vector v possesses contractions which are not connected, then all its contractions are irreducible.*

If the vector v possesses contractions which are not connected, then it follows from Lemma 1 that v possess either two or three contractions. If v

possesses just two contractions, then we infer from Lemma 1, (a) and Postulates VI, VII, that v has the following form:

$v = v' + (a, b, c) + v''$ where $v' + (a)$ and $(c) + v''$ are irreducible vectors and where the sums $a + b$, $b + c$ exist in A, though none of the sums $a + (b + c)$, $(a + b) + c$ exists in A.

If $v' + (a + b, c) + v''$ were reducible, then we would have $v' = w + (t)$ where $t + (a + b)$ would exist in A. But from the existence of the sums $a + b$, $b + c$ and $t + (a + b)$ in A we infer by XI, (b) the existence of at least one of the sums $t + a$ and $(a + b) + c$, which is impossible. Thus the contraction $v' + (a + b, c) + v''$ of v is irreducible; and the irreducibility of the other contraction $v' + (a, b + c) + v''$ of v is derived similarly from XI, (c).

If v possesses just three contractions, then we infer from Lemma 1, (b) and Postulates VI, VII that v has the following form:

$v = w' + (p, q, r, s) + w''$ where $w' + (p)$ and $(s) + w''$ are irreducible vectors and where the sums $p + q$, $q + r$, $r + s$ exist in A, though none of the sums $p + (q + r)$, $(p + q) + r$, $(q + r) + s$, $q + (r + s)$ exists in A.

This clearly contradicts XI, (a).

Thus we have verified the validity of (iii*) and shown that (iii) is a consequence of (ii).

Assume finally the validity of (iii). We first verify (L), and to do this it suffices to prove (L'). This we do by complete induction with respect to the length n of the vector v. It is trivially true, if v has length 1; and thus we may assume that $1 < n$ and that (L') is true for all vectors whose length is less than n. If the vector v of length n is irreducible, then our contention is again trivially true. Thus we assume finally that v is reducible and that v' and v'' are irreducible reductions of v. Consequently there exist contractions w' and w'' such that v' is a reduction of w' and v'' is a reduction of w''. If w' and w'' are both irreducible, then $v' = w'$ and $v'' = w''$ are both of length $n - 1$, as we wanted to show. If w' and w'' are not both irreducible, then they are, by (iii), connected; and there exist consequently contractions $w(0), \cdots, w(k)$ of v such that $w' = w(0)$, $w'' = w(k)$ and such that $w(i)$ and $w(i + 1)$ possess for $0 \leq i < k$ a common contraction $u(i)$. Since $w(i)$ has length $n - 1$, it follows from the inductive hypothesis that all the irreducible reductions of $w(i)$ have the same length $m(i)$. But $u(i)$ has an irreducuible reduction which is necessarily a reduction of both $w(i)$ and of $w(i + 1)$. Consequently $m(i) = m(i + 1)$ for every i, so that in particular

$m(0) = m(k)$. But $m(0)$ is the length of v', and $m(k)$ is the length of v''. Hence v' and v'' have the same length, as we wanted to show. This completes the inductive proof of (L′). Hence (L) too is satisfied in A.

Next we prove:

(S. 1′) *If (a) is a reduction of the vector v, then (a) is the only irreducible reduction of v.*

This is trivially true, if v is of length 1. Thus we may assume that v be of length $n > 1$ and that we have verified (S. 1′) for vectors whose length is less than 1. It follows from (L) that all irreducible reductions of v have length n. Suppose therefore that (b) is some irreducible reduction of v. Then there exist contractions u and w of v such that (a) is a reduction of u and (b) is a reduction of v. If u and w are irreducible, then $v = (r, s)$ is of length 2 and $(a) = (r + s) = (b)$ so that $a = r + s = b$ too. If u and w are not irreducible, then we infer from (iii) the existence of contractions $w(i)$ of v such that $u = w(0)$, $w = w(k)$ and such that $w(i)$ and $w(i + 1)$ possess a common contraction $u(i)$. Since the $w(i)$ are of length $n - 1$, and since all their irreducible reduction have length 1 [as reductions of v], it follows that $w(i)$ has one and only one irreducible reduction (a_i). But (a_i) is a reduction of $u(i)$, as is (a_{i+1}); and so it follows that $a_i = a_{i+1}$. This implies in particular the equality of $a = a_0$ and $a_k = b$, so that the inductive proof of (S. 1′) has been completed. From (S. 1′) one deduces readily (S. 1); and thus we have shown that (i) is a consequence of (iii). This completes the proof of our theorem.

COROLLARY. *If Postulates* I *and* II *are satisfied by the add A, then the associative law* (T) *is a necessary and sufficient condition for the validity of Postulates* VI, VII *and*

XI-. *If $a - b$, $b - c$ and $c - d$ exist in A, then $a - c$ or $b - d$ exists in A.*

Proof. If (T) is satisfied by A, then we deduce the validity of VI, VII from Theorem 1. If furthermore $a - b$, $b - c$, $c - d$ exist in A, then we have

$$(a - b, b - c, c - d) \sim (a, -b, b, -c, c, -d) \sim (a, -d).$$

Since the first of the vectors has length 3 and the last one has length 2, it follows from (T) [invariance of length] that at least one of the sums $(a - b) + (b - c)$ and $(b - c) + (c - d)$ exists in A. But $(a - b) + b = a$ by VI, VII; and thus it follows from VI, VII that $(a - b) + (b - c) = a - c$, if it exists; and similarly one sees that $(b - c) + (c - d) = b - d$, if it exists. Hence XI- is a consequence of (T).

Assume conversely the validity of VI, VII and XI-. Then we are assured of the validity of Postulates I to V [II, §1]. If the sums $a+b$, $b+c$, $c+d$ exist in A, then it follows from I to VII that the differences $(a+b)-b$, $b-(-c)$, $-c-[-(c+d)]$ exist in A; and it follows from XI- that $(a+b)-(-c)=(a+b)+c=a+(b+c)$ or $b-[-(c+d)]=b+(c+d)=(b+c)+d$ exists in A. Thus XI, (a) is true. If the sums $a+b$, $b+c$ and $(b+c)+d$ exist in A, then we infer from Postulates I to VII the existence of the differences $a-(-b)$, $-b-[-(b+c)]$, $-(b+c)-[-((b+c)+d)]$; and it follows from XI- that at least one of the differences $a-[-(b+c)]=a+(b+c)$ or $-b-[-((b+c)+d)]=-b+[(b+c)+d]=c+d$ [because of $-b+(b+c)=c$] exists in A. Thus XI, (c) is true too; and the validity of XI, (b) is verified likewise. Hence VI, VII, XI are true; and (T) is a consequence of Theorem 1.

THEOREM 2. *The following properties of the add A are equivalent.*

(i) *A is the amalgam of its subgroups and satisfies the associative law* (T).

(ii) *A is the amalgam of its subgroups and satisfies the following condition*:

(D) *If X, Y, Z are maximal subgroups of A, then $X \cap Y \leq Y \cap Z$ or $Z \cap Y \leq Y \cap X$.*

(iii) *A satisfies Postulates* I, II, VI, VII *and*

XI*. *If the sums $a+b$, $b+c$, $c+d$ exist in A, then at least one of the sums $a+c$ and $b+d$ exists in A.*

Proof. Assume first the validity of (i). Then A is an amalgam which satisfies the strong associative law; and A is consequently an amalgam in the strict sense [II, §3, Theorem]. Suppose now that X, Y, Z are maximal subgroups of A; and that $X \cap Y \not\leq Y \cap Z$ and $Z \cap Y \not\leq Y \cap X$. Then there exists an element a in $X \cap Y$, but not in $Y \cap Z$. Since a is not in the maximal subgroup Z of A, there exists [II, §2, Corollary 1, (c)] an element c in Z such that the sum $a+c$ does not exist in A. Likewise there exists an element b in $Z \cap Y$ and an element d in X such that the sum $b+d$ does not exist in A. Since d, a are in X, a, b in Y and b, c in Z, we are assured of the existence of the sums $d+a$, $a+b$, $b+c$. Then it follows from Postulate XI, (a) [which is true because of (T) and Theorem 1] that at least one of the sums $d+(a+b)$ and $(a+b)+c$ exists in A. If the first of these sums exists

in A, then we are assured of the existence of the sums $a+d$, $d+(a+b)$, $(a+b)+a$ in the amalgam A in the strict sense; and hence it follows [II, §2, Definition 2, (A. 3)] that $a, d, a+b$ are contained in some subgroup of A. But this subgroup would contain d and b too, which is impossible. If the sum $(a+b)+c$ would exist in A, then we would show likewise that a and c belong to the same subgroup of A, which is impossible too. Thus we are led to a contradiction which proves the validity of Condition (D). Hence (ii) is a consequence of (i).

Assume now the validity of (ii). If X, Y, Z are maximal subgroups of A, then it follows from (D) that the cross cuts $X \cap Y$, $Y \cap Z$, $Z \cap X$ are all contained in one of the subgroups X, Y, Z. Hence A is an amalgam in the strict sense [II, §3, Lemma 2]. Consequently the Postulates I to X are satisfied by A [II, §2]. Assume now the existence of the sums $a+b$, $b+c$, $c+d$ in A. Then there exist maximal subgroups X, Y, Z of A such that a, b are in X, b, c in Y and c, d in Z. Hence b belongs to $X \cap Y$ and c to $Y \cap Z$. If $X \cap Y \leq Y \cap Z$, then b is in Z, as are c and d so that the sum $b+d$ exist in A. If $X \cap Y \not\leq Y \cap Z$, then it follows from (D) that $Z \cap Y \leq Y \cap X$. Hence c belongs to X, as do a and b, so that the sum $a+c$ exists in A. Thus we have verified the validity of XI*; and we have shown that (iii) is a consequence of (ii).

Assume finally the validity of (iii). Then we are assured of the validity of Postulates I to VII and may use the operation $-x$ as usual [II, §1]. We show first the validity of

IX*. *The existence of $a+b$ implies the existence of $a-b$.*

If $a+b$ exists in A, then we are assured of the existence of the sums $a+b$, $b-b\ [=0]$, $-b-a\ [=-(a+b)]$ in A. This implies by XI* the existence of at least one of the sums $a-b$ and $b-a$. But $-(b-a) = -(-a)-b = a-b$, so that $a-b$ exists in either case.

If the sums $x_1-x_2, \cdots, x_{n-1}-x_n, x_n-x_1$ exist in A, and if $3<n$, then we infer from IX* the existence of the sums x_1+x_2, x_2+x_3, x_3+x_4. It follows from XI* that at least one of the sums x_1+x_3 and x_2+x_4 exists in A; and hence it follows from IX* that x_1-x_3 or x_2-x_4 exists in A. Thus we have verified the validity of II, §4, Theorem, (iii) and this shows the validity of the strong associative law.

From the existence of $a-a\ [=0]$ and IX* we infer the existence of $a-(-a)=a+a$. Thus X is satisfied by A. If $a+b$ exists in A, then we deduce from IX* successively the existence of $a-b$, $-(a-b)=b-a$, $b-(-a)=b+a$, so that IX is valid in A. Now it is a consequence of

II, §3, Theorem that A is an amalgam in the strict sense; and this implies in particular the validity of the postulates I to X [II, §2].

If the sums $a+b$, $b+c$, $c+d$ exist in A, then we infer from XI* that at least one of the sums $a+c$ and $b+d$ exists in A. If $a+c$ exists in A, then we are assured of the existence of the sums $a+b$, $b+c$, $a+c$; and it follows from VIII that $(a+b)+c=a+(b+c)$ [by VI, VII] exists in A; and if $b+d$ exists in A, then we see likewise the existence of $(b+c)+d$. Thus XI, (a) holds in A.

If the sums $a+b$, $b+c$ and $(b+c)+d$ exist in A, then we infer from the properties of amalgams the existence of the sums

$$a+b,\ b+(b+c),\ (b+c)+d.$$

Hence it follows from XI* that at least one of the sums $a+(b+c)$ and $b+d$ exists in A. In the latter case we are assured of the existence of the sums $d+b$, $b+(b+c)$, $(b+c)+d$; and it follows from the properties of amalgams in the strict sense [(A. 3)] that $d, b, b+c$ belong to some subgroup of A. This subgroup contains necessarily d, b, c so that the sum $c+d$ exists in A. Thus we have verified the validity of XI, (c); and the validity of XI, (b) is shown likewise. Hence XI is true; and it follows from Theorem 1 that (T) is satisfied by A. We have shown therefore that (i) is a consequence of (iii); and this completes the proof.

Remark. We note that adds A which satisfy the conditions of Theorem 2 satisfy the strong associative law and are amalgams in the strict sense, as was shown during the proof. Thus they belong to a class of adds, investigated in I, §6.

Example. Form the direct sum $H \oplus J \oplus K \oplus L$ of groups H, J, K, L all different from 0. Form the subadd A of $H \oplus J \oplus K \oplus L$ which consists of the elements in $H \oplus J$, $J \oplus K$, $K \oplus L$. It is a consequence of II, §4, Corollary 2 that A is an amalgam in the strict sense which satisfies the strong associative law. But it follows from Theorem 2 that the associative law (T) is not satisfied, since Condition (D) does not hold in A.

2. The similarity problem. A solution of the similarity problem consists in devising a definite method for deciding in a finite number of steps whether two given vectors are similar or not. This problem is often referred to as the identity problem, since similar vectors represent the same element in the derived manifold [Baer (1), §2], and also as the word problem, since the term "word" is often used instead of our "vector."

Every vector possesses irreducible reductions. If the vector v has

length n, then one may obtain each of its irreducible reductions in less than n steps. The procedure may be normalized, if we care to do so, by requiring to make always that contraction which is "farthest left." Likewise it would be possible to obtain all irreducible reductions of v in less than $\frac{1}{2}n(n-1)$ steps. Since a vector and its reductions are similar vectors, it follows that it suffices to solve the similarity problem for irreducible vectors. We are going to offer here a solution of this problem for adds which satisfy Postulates I, II and the associative law (T). This solution is essentially contained in the following proposition.

THEOREM. *If Postulates* I, II *and the associative law* (T) *are satisfied by the add A, then the following conditions are necessary and sufficient for similarity of the irreducible vectors $(a_1, \cdots, a_m)$ and $(b_1, \cdots, b_n)$.*

(a) $m = n$.

(b′) *If* $1 = m = n$, *then* $a_1 = b_1$.

(b″) *If* $1 < m = n$, *then there exist elements* $c_2, \cdots, c_n$ *in* A *such that*

$$c_2 = -b_1 + a_1,$$
$$c_i = -b_{i-1} + (c_{i-1} + a_{i-1}) = (-b_{i-1} + c_{i-1}) + a_{i-1} \text{ for } 2 < i \leq n,$$
$$c_n = b_n - a_n.$$

Proof. It is a consequence of §1, Theorem 1 that Postulates VI, VII are satisfied by A. Consequently Postulates I to VII are satisfied by A; and we may use the operation $-x$ [II, §1].

Assume now the validity of our conditions. If $m = n = 1$, then our vectors are equal [and consequently similar]. If $1 < m = n$, then it follows from our conditions [and $y = -x + (x + y)$] that

$$a_1 = b_1 + c_2,\ c_{i-1} + a_{i-1} = b_{i-1} + c_i \text{ for } 2 < i \leq n,\ b_n = c_n + a_n.$$

This implies

$$\begin{aligned}(a_1, \cdots, a_m) &\sim (a_1, -c_2, c_2, a_2, \cdots, a_{i-1}, -c_i, c_i, a_i, \cdots, -c_n, c_n, a_n)\\ &\sim (a_1, -c_2, c_2 + a_2, \cdots, c_{i-1} + a_{i-1}, -c_i, c_i + a_i, \cdots, -c_n, c_n + a_n)\\ &= (b_1 + c_2, -c_2, b_2 + c_3, \cdots, b_{i-1} + c_i, -c_i, b_i + c_{i+1}, \cdots, -c_n, b_n)\\ &\sim (b_1, c_2, -c_2, b_2, c_3, \cdots, b_{i-1}, c_i, -c_i, b_i, c_{i+1}, \cdots, -c_n, b_n)\\ &\sim (b_1, \cdots, b_n);\end{aligned}$$

and thus we have shown the sufficiency of our conditions.

Suppose conversely the similarity of the irreducible vectors $(a_1, \cdots, a_m)$

and $(b_1, \cdots, b_n)$. Then $m = n$ is a consequence of the invariance property (L). If $1 = m = n$, then it follows from (S. 1) that $a_1 = b_1$. Thus we have shown the necessity of our condition (b) in case the similar irreducible vectors are of length 1. Hence we asume now that the irreducible similar vectors are of equal length $n > 1$ and that the validity of (b) has been verified for similar irreducible vectors of length less than n. From the similarity of $(a_1, \cdots, a_n)$ and $(b_1, \cdots, b_n)$ we deduce now that

$$(a_1, \cdots, a_{n-1}) \sim (a_1, \cdots, a_{n-1}, a_n, -a_n) = (a_1, \cdots, a_n) + (-a_n)$$
$$\sim (b_1, \cdots, b_n) + (-a_n) = (b_1, \cdots, b_n, -a_n).$$

Since the first of these vectors has length $n-1$ and the last one has length $n+1$, it follows from (L) that the last vector is reducible. Since $(b_1, \cdots, b_n)$ is irreducible, this shows the existence of

$$c_n = b_n - a_n.$$

Hence $(a_1, \cdots, a_{n-1}) \sim (b_1, \cdots, b_{n-1}, c_n)$; and it follows as before that the sum $b_{n-1} + c_n$ exists in A. Consequently we have

$$(a_1, \cdots, a_{n-1}) \sim (b_1, \cdots, b_{n-2}, b_{n-1} + c_n).$$

The first of these vectors of length $n-1$ is irreducible by hypothesis and thus we may deduce the irreducibility of the second vector from the invariance rule (L). Upon these similar irreducible vectors of length $n-1$ the inductive hypothesis may be applied. Consequently either $n-1=1$ and $a_1 = b_1 + c_2$ or $c_2 = -b_1 + (b_1 + c_2) = -b_1 + a_1$, or else $1 < n-1$ and there exist elements $c_2, \cdots, c_{n-1}$ in A such that $c_2 = -b_1 + a_1$,

$$c_i = -b_{i-1} + (c_{i-1} + a_{i-1}) = (-b_{i-1} + c_{i-1}) + a_{i-1} \text{ for } 2 < i < n,$$

$c_{n-1} = (b_{n-1} + c_n) - a_{n-1}$. From the last of these equations we deduce first that

$$b_{n-1} + c_n = [(b_{n-1} + c_n) - a_{n-1}] + a_{n-1} = c_{n-1} + a_{n-1},$$
$$c_n = -b_{n-1} + (b_{n-1} + c_n) = -b_{n-1} + (c_{n-1} + a_{n-1}).$$

We have furthermore

$$(b_1, \cdots, b_{n-1}) \sim (b_1, \cdots, b_{n-1}, c_n, -c_n)$$
$$\sim (b_1, \cdots, b_{n-2}, b_{n-1} + c_n) + (-c_n)$$
$$\sim (a_1, \cdots, a_{n-1}) + (-c_n) = (a_1, \cdots, a_{n-1}, -c_n);$$

and now we deduce as usual from (L) the existence of $a_{n-1} - c_n$. This

implies the existence of $-(a_{n-1}-c_n)=c_n-a_{n-1}$; and consequently it follows from Posulates VI, VII that

$$c_{n-1}=(b_{n-1}+c_n)-a_{n-1}=b_{n-1}+(c_n-a_{n-1}),$$

$$c_n-a_{n-1}=-b_{n-1}+[b_{n-1}+(c_n-a_{n-1})]=-b_{n-1}+c_{n-1},$$

$$c_n=(c_n-a_{n-1})+a_{n-1}=(-b_{n-1}+c_{n-1})+a_{n-1};$$

and this clearly completes the inductive proof of (b).

Remark 1. This theorem contains a solution of the similarity problem, since it gives a definite procedure for finding elements $c_2,\cdots,c_n$—if they cannot be determined in this fashion, then the vectors are not similar—and since it subjects the last element c_n to a further test.

Remark 2. In case A happens to be an amalgam satisfying (T), then one may prove easily the following fact: If C is a class of irreducible similar vectors over A, then all the vectors in C have the same length $n(C)$; and there exists a subgroup $S(i)$ of A which contains all the i-th coordinates of vectors in C.

3. Equality of similar irreducible vectors. If similar irreducible vectors are necessarily equal, then the similarity problem, discussed in §2, has a trivial solution. We are going to characterize these adds.

THEOREM 1. *The following properties of the add A are equivalent.*

(i) *Similar irreducible vectors are equal.*

(ii) *A satisfies Postulates* VI *and*

XII. *If the sums $a+b$, $b+c$ exist in A, though at least one of the sums $a+(b+c)$ and $(a+b)+c$ does not exist, then $a=a+b$ and $c=b+c$.*

(iii) *All contractions of the vector v are connected.*

Proof. Assume first that similar irreducible vectors are equal. If the sums $a+b$ and $b+c$ exist in A, then we have

$$(a+b,c)\sim(a,b,c)\sim(a,b+c).$$

Since the first and last of these vectors are of length 2, it follows that they are either both reducible or both irreducible. If they are both reducible, then we have $((a+b)+c)\sim(a+(b+c))$; and this implies $(a+b)+c=a+(b+c)$. If they are both irreducible, then they are equal, and hence $a+b=a$, $c=b+c$. This shows that (ii) is a consequence of (i).

Assume next the validity of (ii). Then one deduces readily the validity of Postulates VI, VII, XI; and it follows from §1, Theorem 1 that the contractions v' and v'' of the vector v are both irreducible, if they are not connected. It follows then from §1, Lemma that

$$v = u + (a, b, c) + w, \quad v' = u + (a + b, c) + w, \quad v'' = u + (a, b + c) + w,$$

where neither of the sums $(a + b) + c$ and $a + (b + c)$ exists in A. Apply XII to show that $a = a + b$, $c = b + c$ and that therefore $v' = v''$. This is a contradiction, since we assumed that v' and v'' are not connected. Hence we have shown that (iii) is a consequence of (ii).

Assume finally the validity of (iii). Then we are going to show first that

(E′) *every vector possesses one and only one irreducible reduction.*

This is certainly true for vectors v of length 1. Hence we assume that v is a vector of length $n > 1$; and that vectors of length less than n possess one and only one irreducible reduction. There is nothing to prove, if v is irreducible. Hence assume that v be a reducible vector of length n; and that w', w'' are irreducible reductions of v. Then there exist contractions v' and v'' of v such that w' and w'' are reductions of v' and v'' respectively. Then either $v' = v''$, or there exist contractions $v(1), \cdots, v(k)$ of v such that $v' = v(1)$, $v(k) = v''$ and such that $v(i)$ and $v(i + 1)$ possess a common contraction. It follows from the inductive hypothesis that v', v'' and the $v(i)$—as vectors of length $n - 1$—possess one and only one irreducible reduction. This implies $w' = w''$ in case $v' = v''$; and if $v' \neq v''$, then $v(i)$ and $v(i + 1)$ have the same irreducible reduction, since they possess a common contraction and therefore a common irreducible reduction. Hence it follows also in the later case that $v' = v(1)$ and $v'' = v(k)$ possess the same irreducible reduction, namely $w' = w''$. This completes the inductive proof of (E′); and from (E′) we deduce (i) by the customary arguments. This completes the proof.

THEOREM 2. *The following properties of the add A are equivalent.*

(i) *A is an amalgam such that similar irreducible vectors are equal.*

(ii) *A satisfies Postulates* I, II, VI, VII *and*

XII.* *If the sums $a + b$, $b + c$ exist in A, and if $b \neq 0$, then the sum $a + c$ exists in A.*

(iii) *A is an amalgam such that distinct maximal subgroups of A have only 0 in common.*

Proof. If (i) is true, then the strong associative law is satisfied by A. Consequently A is an amalgam in the strict sense and satisfies Postulates I to X [II, §3, Theorem & II, §2]. If the sums $a+b$, $b+c$ exist in A, and if $b \neq 0$, then the non-existence of one of the sums $a+(b+c)$ and $(a+b)+c$ would imply [by Theorem 1, XII] that $a = a+b$ and $c = b+c$; and this would imply $b = 0$, since a and b as well as b and c belong to some subgroup of A. Consequently the sums $a+(b+c)$, $(a+b)+c$ exist in A; and this implies as usual that a, b, c belong to some subgroup of A. Hence $a+c$ exists in A. Thus (ii) is a consequence of (i).

If (ii) is true, then one verifies readily the validity of Postulates I, II, VI, VII and XI*; and it follows from §1, Theorem 2 that A is an amalgam in the strict sense which satisfies the associative law (T). If X and Y are distinct maximal subgroups of A, then there exists an element x in X and and an element y in Y such that the sum $x+y$ does not exist in A [II, §2, Corollary 1, (c)]. If z belong to the cross cut of X and Y, then the sums $x+z$, $z+y$ exist in A, though $x+y$ does not exist in A. Hence we deduce from XII* that $z = 0$ or $X \cap Y = 0$. This shows the validity of (iii) so that (iii) is a consequence of (ii).

If finally (iii) is satisfied, and if the sums $a+b$, $b+c$ exist in A, though the sums $a+(b+c)$ and $(a+b)+c$ do not both exist in A, then there does not exist a subgroup of A which contains a, b, c. Consequently there are distinct maximal subgroups X, Y such that a, b are in X and b, c in Y. Then b is in $X \cap Y = 0$ so that $b = 0$. This shows the validity of Postulate XII; and we deduce from Theorem 1 that similar irreducible vectors are equal. Hence (i) is a consequence of (iii). This completes the proof.

Remark. This theorem shows that the free sums of groups are characterized among the free sums with amalgamations by the property that similar irreducible vectors are equal.

4. The transformation problem. Throughout this section we shall be assured of the validity of the Postulates I to V; as a matter of fact we shall mostly, though not always, assume that the adds under consideration are amalgams in the strict sense [so that all the Postulates I to X are valid]. Thus there exists to every element x in the add A a uniquely determined element $-x$ such that $x-x = -x+x = 0$; and we shall use the notation: $-(v_1, \cdots, v_n) = (-v_n, \cdots, -v_1)$.

DEFINITION 1. *The vectors v and w over A are equivalent, if there exists a vector t over A such that $t+v \sim w+t$.*

Since the defining similarity may also be written in the form $v \sim -t + w + t$, this amounts to saying that equivalent vectors are "transforms" of each other. Noting this, it follows that similarity implies equivalence, and that equivalence is symmetric, reflexive, transitive and is in agreement with generally accepted terminology.

One proves easily that the vectors v and w are equivalent if, and only if, there exists a vector u such that $v \sim u$ and w arises from u by a cyclic permutation of the coordinates. Thus we are led to the following concepts.

DEFINITION 2. *The vector v is a cyclic reduction of the vector w, if ther exist vectors $t(0), \cdots, t(k)$ such that $w = t(0)$, $t(k) = v$, and such that $t(i+1)$ is either a reduction of $t(i)$ or else is obtained from $t(i)$ by a cyclic permutation of the coordinates of $t(i)$.*

It is clear that every reduction is also a cyclic reduction and that cyclic reductions of v are equivalent to v.

DEFINITION 3. *If all the vectors arising from v by cyclic permutation of coordinates are irreducible, then v is cyclically irreducible.*

Thus $(v_1, \cdots, v_n)$ is cyclically irreducible, if either $n = 1$ or else none of the sums $v_i + v_{i+1}$ for $0 < i < n$ nor the sum $v_n + v_1$ exists in A.

It is easy to see that every vector possesses cyclic reductions which are cyclically irreducible; and that cyclically irreducible vectors do not possess "proper" cyclic reductions.

DEFINITION 4. *The elements a and b are equivalent over A, if there exists a vector v of non-negative length such that v and $-v$ are irreducible vectors and such that (a) is a reduction of $v + (b) - v$.*

If the strong associative law holds in A, then this concept of equivalence is clearly symmetric, reflexive, and transitive. Whether the strong associative law is needed for the validity of this statement, is an open question.

If we substitute in our various "associative laws" equivalence for similarity of vectors and equivalence for equality of elements, cyclic reduction for reduction and cyclical irreducibility for irreducibility, then we obtain a new kind of associative laws. It is our object to prove the rather surprising fact that, at least for amalgams, these new associative laws are equivalent to the original laws they arose from.

If we apply the procedure just outlined to the strong associative law, then we obtain the following property.

THE ASSOCIATIVE LAW (S_t) consists of the following two properties:

(S_t. 1) *If the vectors* (a) *and* (b) *are equivalent, then the elements* a *and* b *are equivalent over* A.

(S_t. 2) *If the cyclically irreducible vectors* v *and* w *are equivalent, and if one of them has length* 1, *then both have length* 1.

The second property will usually be used in the following form: If (a) and v are equivalent, and if the length of v is greater than 1, then v is cyclically reducible.

THEOREM 1. *The strong associative law and the associative law* (S_t) *are equivalent properties of amalgams.*

This theorem will be a consequence of two slightly stronger propositions.

THEOREM 1a. *The associative law* (S_t) *is a consequence of the strong associative law and the Postulates* I, II, IX.

Proof. It is a consequence of the strong associative law and Postulates I, II that Postulates I to V are satisfied by A [see II, §4, Theorem]. If (a) and (b) are equivalent vectors over A, then there exists a vector v of shortest length such that $(a) \sim -v + (b) + v$. It is clear that v and $-v$ are irreducible vectors; and it follows from the strong associative law that (a) is a reduction of $-v + (b) + v$. Hence a and b are equivalent elements over A [Definition 4]; and thus (S_t. 1) is satisfied by A.

Suppose now that the vector $v = (v_1, \cdots, v_m)$ is cyclically irreducible, has length m greater than 1, and is equivalent to a vector of length 1. Then there exists a vector w of minimal length n such that $w + v - w$ is similar to a vector of length 1. One verifies readily that w and $-w$ are irreducible vectors; and one deduces from the strong associative law that $w + v - w$ possesses a reduction of the form (a). It is now easy to show that $w + v - w = p + q$ where the vectors p and q possess reductions (b) and (c) respectively such that $a = b + c$. If the length of p would not exceed n, then we would have $w = p + p'$. Since w is irreducible, so is p. But (b) is a reduction of p so that $p = (b)$. From the existence of the sum $b + c = a$ and Postulate IX we infer the existence of the sum $c + b = d$. Hence (d) is a reduction of $q + p$. Now we see that

$$(d) \sim q + p = p' + v - p' - p + p \sim p' + v - p';$$

and this is impossible, since p' has length at most $n - 1$, and since w was selected as a vector of minimal length transforming v into a vector of length 1. Thus we have shown that the length of p is greater than n; and one verifies

likewise that the length of q is greater than n. Consequently there exists an integer i such that

$$0 < i < m, \quad p = w + (v_1, \cdots, v_i), \quad q = (v_{i+1}, \cdots, v_m) - w.$$

From the existence of the sum $a = b + c$ and Postulate IX we infer again the existence of the sum $c + b = d$. Since (b) and (c) are reductions of p and q respectively, we find that (d) is a reduction of $q + p$. Consequently we have

$$(d) \sim q + p = (v_{i+1}, \cdots, v_m) - w + w + (v_1, \cdots, v_i)$$
$$\sim (v_{i+1}, \cdots, v_m, v_1, \cdots, v_i) = v'.$$

It follows from the strong associative law that (d) is a reduction of v'. But v' arises from v by cyclic permutation of coordinates; and this contradicts the assumed cyclical irreducibility of v. Thus we have been led to a contradiction which proves the validity of $(S_t. 2)$. This completes the proof.

Remark. It appears probable that Postulate IX is not necessary for the validity of (S_t), though it may be indispensable.

THEOREM 1b. *The strong associative law is a consequence of the associative law* (S_t) *and the Postulates* I *to* V.

Proof. It is a consequence of II, §**4**, Theorem that it suffices to prove the validity of the following proposition:

(*) *If* $(0) \sim v$, *then* (0) *is a reduction of* v.

This we prove by complete induction with respect to the length n of the vector v. If firstly $v = (a)$ is of length 1, then we infer from $(0) \sim (a)$ the equivalence of (0) and (a); and we infer from Property $(S_t. 1)$ the equivalence of the elements 0 and a in A. Hence there exists a vector t such that t and $-t$ are irreducible vectors and such that (0) is a reduction of $t + (a) - t$. It follows from II, §**4**, Lemma 2 that (0) is likewise a reduction of $-t + t + (a)$. We infer from II, 4, Lemma 1, (b) the existence of a reduction (b) of $-t + t$ such that $b + a = 0$. But t and $-t$ are irreducible vectors; and thus (0) is the only reduction of $-t + t$. Hence $b = 0$, and consequently $a = 0$. This proves (*) for vectors v of length 1.

Now we assume that $(0) \sim w$, that w is a vector of length $n + 1$ with $0 < n$, and that we have verified (*) for all vectors v whose length does not exceed n. Similarity implies equivalence; and it follows from Property $(S_t. 2)$ that w is cyclically reducible. We distinguish two cases.

Case 1. w is reducible. Then there exists a contraction w' of w. Since

clearly $(0) \sim w \sim w'$, and since w' has length n, we may apply the inductive hypothesis which assures us that (0) is a reduction of w' and hence of w.

Case 2. w is irreducible. Then it follows from the cyclic reducibility of w that $w = (r) + s + (t)$ and that the sum $t + r$ exists in A. We have $(t, r) + s \sim (t) + w - (t) \sim (0)$. Since $(t, r) + s$ is reducible, it follows from Case 1 that (0) is a reduction of $(t, r) + s$. But then it follows from II, §4, Lemma 2 that (0) is likewise a reduction of the vector w which arises from $(t, r) + s$ by a cyclic permutation of coordinates. This completes the induction, the proof of (*) and of Theorem 1b, as has been pointed out before.

That Theorem 1 is an immediate consequence of Theorems 1a and 1b, is fairly obvious, since Postulates I to V and IX are satisfied by all amalgams.

The Associative Law (T_t) consists of the Property $(S_t. 1)$ and of the following invariance property:

(L_t) *Equivalent cyclically irreducible vectors have the same length.*

It is clear that (L_t) implies $(S_t. 2)$. Hence (S_t) is a consequence of (T_t). If, furthermore, Postulates I to V are satisfied, then it follows from Theorem 1b that the strong associative law is likewise satisfied by A.

The following notation will prove convenient, since we are going to restrict our discussion to amalgams.

Definition 5. *The subset S of the add A is co-group, if there exists a subgroup T of A such that $S \leq T$.*

Co-group subsets are clearly contained in maximal subgroups, since every subgroup is contained in a maximal subgroup.

Theorem 2. *The associative law* (T) *and the associative law* (T_t) *are equivalent properties of amalgams.*

Proof. We have noted before that the strong associative law is always a consequence of (T); and that in the presence of Postulates I to V [which are satisfied in all amalgams] the strong associative law is likewise a consequence of (T_t). Thus we may assume throughout the proof that the strong associative law is satisfied by the amalgam A; and that therefore A is an amalgam in the strict sense of the word [II, §3, Theorem]. Thus all the postulates I to X are satisfied by A [II, §2] and also the property:

The elements x, y, z are co-group if, and only if, the sums $x + y$, $y + z$, $z + x$ exist in A.

We note furthermore the equivalence of the following three properties of elements a, b in A:

$$a, b \text{ are co-group}; \ a + b \text{ exists in } A; \ a - b \text{ exists in } A.$$

A. Assume the validity of (T_t). Consider elements a, b, c, d in A whose differences $a - b$, $b - c$, $c - d$ exist in A. Then we have the following vector similarities:

$$(a - b, b - c, c - d) \sim (a, -b, b, -c, c, -d) \sim (a, -d).$$

Thus we have shown the similarity, and hence the equivalence, of a vector of length 3 and a vector of length 2. It follows from (L_t) that the vector of length 3 is not cyclically irreducible. Hence at least one of the sums $(a - b) + (b - c)$, $(b - c) + (c - d)$ and $(c - d) + (a - b)$ exists in A.

Case 1. $(a - b) + (b - c)$ exists in A. Then we have the vector similarities:

$$((a - b) + (b - c)) \sim (a, -b, b, -c) \sim (a, -c);$$

and we infer from the strong associative law the existence of the difference $a - c$.

Case 2. $(b - c) + (c - d)$ exists in A. Then one deduces [as in Case 1] the existence of $b - d$.

Case 3. $(c - d) + (a - b)$ exists in A. Then there exists $(a - b) + (c - d)$ and we have

$$((a - b) + (c - d)) \sim (a, -b, c, -d) \sim (a, -b + c, -d).$$

[Note that the existence of $-b + c$ is a consequence of the existence of $c - b$.] It follows from the strong associative law that at least one of the sums $a + (-b + c)$ and $(-b + c) - d$ exists in A. If $a + (-b + c)$ exists in A, then the pairs $a, -b + c$ and $-b + c, -b$ and $a, -b$ are pairs of co-group elements in A. But A is an amalgam in the strict sense so that $a, -b + c, -b$ are co-group. Hence a, b, c are co-group so that $a - c$ exists in A. If $(-b + c) - d$ exists in A, then we show likewise the existence of $b - d$ in A.

Thus we have shown that in every case at least one of the differences $a - c$, $b - d$ exists in A. The validity of (T) is therefore a consequence of §1. Corollary.

B. Assume conversely the validity of (T). Then one infers immediately from §1, Theorem 2, (iii) and the fact that A is an amalgam in the strict sense the validity of the following proposition.

(B. 1) *If a, b and b, c and c, d are three pairs of co-group elements in A, then at least one the triplets a, b, c and b, c, d is co-group.*

Next we prove the following fundamental proposition.

(B. 2) *Suppose that v and w are similar irreducible vectors. Then v is cyclically irreducible if, and only if, w is cyclically irreducible.*

It is a consequence of the invariance of length that v and w are vectors of equal length n. Because of the symmetry of our hypotheses it suffices to show that cyclical reducibility of w implies cyclical reducibility of v. The irreducible, cyclically reducible vector w has the form $w = (w_1, \cdots, w_n)$, where none of the sums $w_i + w_{i+1}$ with $0 < i < n$ exists in A, whereas the sum $w_n + w_1$ exists in A, so that in particular $2 < n$. Let $v = (v_1, \cdots, v_n)$. Then we infer from $v \sim w$ and §2, Theorem the existence of $c = w_n - v_n$, $d = w_1 - v_1$ and the fact that c, w_{n-1}, v_{n-1} and d, w_2, v_2 are co-group triplets. We have furthermore the following vector similarities:

$$(w_n + w_1, w_2, \cdots, w_{n-2}, w_{n-1} + c) \sim (w_n, w_1, \cdots, w_{n-1}, c)$$
$$\sim (w_n) + w + (-v_n) \sim (w_n) + v + (-v_n) \sim (w_n, v_1, \cdots, v_{n-1}).$$

Since the first of these vectors has length $n - 1$, whereas the length of the last one is n, it follows from the invariance of length that the last of these vectors is reducible; and this implies because of the irreducibility of v the existence of the sum $w_n + v_1$. Thus we have seen that the three pairs w_n, v_1 and v_1, w_1 and w_1, w_n are co-group. But A is an amalgam in the strict sense; and thus it follows that w_n, w_1, v_1 are co-group. Interchanging 1 and n it follows that w_1, w_n, v_n are co-group too. Hence there exist maximal subgroups X, Y, Z of A such that v_n, w_n, w_1 are in X, v_1, w_1, w_n in Y and d, w_2, v_2 in Z. If $X \cap Y \leq Y \cap Z$, then w_1, w_n would be in Z. But w_2 is in Z and w is irreducible. Hence $X \cap Y \not\leq Y \cap Z$; and it follows from §1, Theorem 2 that $Z \cap Y \leq Y \cap X$. But $d = w_1 - v_1$ belongs to $Z \cap Y$ and hence to X. Hence $w_1 - v_1, v_n, w_1$ are in X; and this implies the existence of $v_n + v_1$ so that v is cyclically reducible too, as we claimed.

(B. 3) *If w is a cyclically irreducible vector of length $m > 1$, if $z = (z_1, \cdots, z_n)$ is an irreducible vector of positive length, and if $z + w - z$ is similar to a cyclically irreducible vector, then $(z_n) + w - (z_n)$ is similar to a cyclically irreducible vector of length m.*

As a matter of fact we are going to prove slightly more, namely that $(z_n) + w + (-z_n)$ possesses a cyclically irreducible reduction of length m. The proof of this fundamental proposition will be effected in a number of steps.

(a) *If the sums $a + b$, $(a + b) + c$, $[(a + b) + c] + d$ exist in A, then at least one of the triplets a, b, c and $a + b, c, d$ is co-group.*

It follows from our hypotheses that the pairs a, $a + b$ and $a + b$, $(a + b) + c$ and $(a + b) + c, d$ are co-group; and hence it follows from (B. 1) that at least one of the triplets a, $a + b$, $(a + b) + c$ and $a + b$, $(a + b) + c, d$ is co-group. In the first case a, b, c are co-group; and in the second case $a + b, c, d$ are co-group.

(b) *Suppose that $w = (w_1, \cdots, w_n)$ is a cyclically irreducible vector and that $1 < n$.*

(b. 1) *If the sums $z + w_1$, $(z + w_1) + w_2$ exist in A, then $([z + w_1] + w_2, w_3, \cdots, w_n, -z)$ is a cyclically irreducible vector.*

(b. 2) *If the sums $w_n - z$, $w_{n-1} + (w_n - z)$ exist in A, then $(z, w_1, \cdots, w_{n-2}, w_{n-1} + [w_n - z])$ is a cyclically irreducible vector.*

(b. 3) *If the sums $z + w_1, w_n - z$ exist in A, then $(z + w_1, w_2, \cdots, w_{n-1}, w_n - z)$ is a cyclically irreducible vector.*

If the sums $z + w_1$, $(z + w_1) + w_2$ exist in A, then the sum $[(z + w_1) + w_2] + w_3$ cannot exist in A [this is vacuous, if $n = 2$], since it would otherwise follow from (a) that one of the triplets z, w_1, w_2 and $z + w_1, w_2, w_3$ is co-group, which would contradict the irreducibility of w. Furthermore neither $-z + [(z + w_1) + w_2]$ nor $[(z + w_1) + w_2] - z$ can exist in A, since otherwise both these sums would exist in A, and since this would imply by (a) that at least one of the triplets z, w_1, w_2 and $z + w, w_2, -z$ is co-group. But in either case $w_1 + w_2$ would exist, which is impossible. This proves already (b. 1) in case $n = 2$. To finish the proof of (b. 1) we have to show that $w_n - z$ cannot exist in case $2 < n$. But if $w_n - z$ existed in A, then the pairs w_n, z and $z, z + w_1$ and $z + w_1, w_2$ would be co-group. This would

imply by (B. 1) that at least one of the triplets $w_n, z, z + w_1$ and $z, z + w_1, w_2$ would be co-group. In the first case $w_n + w_1$ would exist in A; and in the second case $w_1 + w_2$. But this contradicts the cyclical irreducibility of w. This completes the proof of (b. 1); and the validity of (b. 2) is shown likewise [or may be inferred by reasons of symmetry].

To prove (b. 3) assume the existence of $z + w_1$ and $w_n - z$. Then it follows from (b. 1) that $(z + w_1) + w_2$ does not exist in A; and it follows from (b. 2) that $w_{n-1} + (w_n - z)$ does not exist in A. The existence of $(w_n - z) + (z + w_1)$ would imply finally that z, w_1, w_n are co-group, which is impossible, since w is cyclically irreducible; and this completes the proof of (b).

Now we are ready to prove (B. 3). Since z is a vector of positive length, $z + w - z$ is not cyclically irreducible. But by hypothesis $z + w - z$ is similar to a cyclically irreducible vector. Hence it follows from (B. 2) that $z + w - z$ is reducible. Since w and z, and consequently also $-z$, are irreducible vectors, it follows that at least one of the sums $z_n + w_1$ and $w_m - z_n$ exists in A. [Here we assume that $w = (w_1, \cdots, w_m)$.] If both $z_n + w_1$ and $w_m - z_n$ exist in A, then it follows from (b. 3) that $(z_n + w_1, w_2, \cdots, w_{m-1}, w_m - z_n)$ is a cyclically irreducible reduction of $(z_n) + w + (-z_n)$. If exactly one of the two sums exists in A, then we may assume wtihout loss in generality that $z_n + w_1$ exists in A, whereas $w_m - z_n$ does not exist in A. If $(z_n + w_1) + w_2$ exists in A, then it follows from (b. 1) that $([z_n + w_1] + w_2, w_3, \cdots, w_m, -z_n)$ is a cyclically irreducible reduction of $(z_n) + w + (-z_n)$. Thus our proof of (B. 3) will be complete, once we have shown that it is impossible to assume that $z_n + w_1$ exists in A, whereas neither $(z_n + w_1) + w_2$ nor $w_m - z_n$ exists in A. If these assumptions were satisfied, then the vector

$$t = (z_n + w_1, w_2, \cdots, w_m, -z_n, -z_{n-1}, \cdots, -z_1)$$

would be irreducible, since w and $-z$ are irreducible vectors. We note the similarity

$$z + w - z \sim (z_1, \cdots, z_{n-1}) + t$$

so that both these vectors are similar to a cyclically irreducible vector.

If $n = 1$, then we infer from (B. 2) that t is cyclically irreducible which contradicts, however, the fact that $-z_1$ and $z_1 + w_1$ are co-group.

If $n = 2$, then it is clear that $(z_1) + t$ is not cyclically irreducible. Since this vector is similar to a cyclically irreducible vector, it follows from

(B. 2) that $(z_1) + t$ is reducible. Hence we may infer from the irreducibility of t the existence of $z_1 + (z_2 + w_1)$. Consequently

$$t^0 = (z_1 + (z_2 + w_1), w_2, \cdots, w_m, -z_2, -z_1) \sim (z_1) + t \sim z + w - z$$

is similar to a cyclically irreducible vector, though it is clear that t^0 itself is not cyclically irreducible, since $-z_1$ and $z_1 + (z_2 + w_1)$ are co-group. It follows from (B. 2) that t^0 is reducible; and this implies that w_2 and $z_1 + (z_2 + w_1)$ are co-group, since t is irreducible. If we interchange right and left in (a)—as we may—then this implies that at least one of the triplets $w_2, z_1, z_2 + w_1$ and z_1, z_2, w_1 is co-group. But this is impossible, since we assumed the irreducibility of z and the non-existence of the sum $(z_2 + w_1) + w_2$.

Thus we may assume finally that $2 < n$. Then we see exactly as in the preceding paragraph of the proof that $z_{n-1} + (z_n + w_1)$ exists in A. Consequently

$$t' = (z_1, \cdots, z_{n-2}, z_{n-1} + (z_n + w_1), w_2, \cdots, w_m) - z \sim z + w - z$$

is similar to a cyclically irreducible vector, though t' is clearly not cyclically irreducible itself. Hence it follows from (B. 2) that t' is reducible; and it follows from the irreducibility of z and w that at least one of the sums

$$z_{n-2} + [z_{n-1} + (z_n + w_1)] \text{ and } [z_{n-1} + (z_n + w_1)] + w_2$$

exists in A. As before, we may now deduce from (a) that at least one of the following three triplets is co-group:

$$z_{n-2}, z_{n-1}, z_n + w_1 \text{ and } z_{n-1}, z_n, w_1 \text{ and } w_2, z_{n-1}, z_n + w_1;$$

and this contradicts the assumed irreducibility of the vectors z and t. Thus we have been led to contradictions in all cases; and this completes the proof of (B. 3).

We are now ready to deduce the associative law (T_t) from (T). We note first that (T) implies the strong associative law; and this implies [by Theorem 1] the associative law (S_t). Thus Property (S_t. 1) is satisfied by A. It follows furthermore from (S_t. 2) that if one of two equivalent, cyclically irreducible vectors has length 1, then both have length 1. Suppose now that v and w are equivalent and cyclically irreducible vectors neither of which has length 1. Then there exists an irreducible vector z of length n such that $z + w - z \sim v$. If z has length 0, then w and v are similar irreducible vectors; and the equality of their length is a consequence of (T). If z has positive length, then $z = z' + (s)$; and it follows from (B. 3) that $(s) + w - (s)$ is similar to a cyclically irreducible vector w' whose length

equals the length of w. But then $v \sim z' + w' - z'$ where z' is an irreducible vector of length $n-1$; and where w' is a cyclically irreducible vector of the same length m as w. By an obvious induction one proves now that v, w' and w have the same length m. Thus (L_t) is true; and this completes the proof of Theorem 2.

Remark. The lemma (b) which we deduced during the proof of the preceding theorem contains a sort of procedure for deriving equivalent cyclically irreducible vectors from a given cyclically irreducible vector; but it does not seem possible to derive from it a straightforward method which could be used to determine whether or not two given cyclically irreducible vectors are equivalent.

THEOREM 3. *Similarity of irreducible vectors over the amalgam A implies their equality if, and only if, A has the following properties:*

(1) *Equivalence of vectors (a) and (b) implies that a and b are conjugate elements in some subgroup of A.*

(2) *Equivalence of the cyclically irreducible vectors v and w which are not both of length 1 implies that v arises from w by a cyclic permutation of the coordinates.*

The proof of this theorem will be omitted, since it is not very difficult. It is best based on §3, Theorem 2, (iii).

UNIVERSITY OF ILLINOIS,
URBANA, ILLINOIS.

BIBLIOGRAPHY.

Reinhold Baer.

(1) "Free sums of groups and their generalizations. An analysis of the associative law," *American Journal of Mathematics*, vol. 71 (1949), pp. 706-742.

(2) "Free sums of groups and their generalizations II," *ibid.*, vol. 72 (1950), pp. 625-646.

GENERALIZED FREE SUMS OF CYCLICAL GROUPS.*

By HANNA NEUMANN.

1. Introduction. The well-known clover knot group (i. e. the fundamental group of the residual space in S_3 of the clover knot) is generated by two elements a and b satisfying the relation $a^2 = b^3$. More generally, any torus knot group is generated by two elements a and b with the defining relation $a^m = b^n$. In such a group both a and b are of infinite order, and the intersection of the two cycles generated by a and b respectively is exactly the cycle generated by the common power $a^m = b^n$, no more. Thus, in Schreier's terminology, the torus knot group is the free product of two infinite cycles with one amalgamated subgroup.

We consider the following generalization: A group is generated by elements $a_1, a_2, \cdots$, finite or infinite in number, with the defining relations

1.1. $\qquad a_i^{m_i} = 1 \qquad (i = 1, 2, \cdots)$,[1]

1.2. $\qquad a_i^{e_{ij}} = a_j^{e_{ji}} \qquad (i \neq j; i, j = 1, 2, \cdots)$.

This group may, or may not, have the following property:

1.3. *All generators a_i are of exact order m_i, and any two generators a_i and a_j have exactly the power $a_i^{e_{ij}} = a_j^{e_{ji}}$ in common, no more.*

Examples of such groups not having this property are easily made. In fact there are some fairly obvious conditions for the numbers e_{ij} which are necessary for the group to satisfy 1.3. They merely express the fact that the relations 1.2 are 'compatible' (in a sense to be precisely defined in **2**) with the relations 1.1 and with each other.

It is the main result of this note that these conditions are also sufficient.

In fact, we shall show more: The compatibility conditions ensure that the group has property 1.3 even if the further relations $a_i a_j = a_j a_i$ are added to 1.1 and 1.2, i. e. if the group is made abelian.

Although the theorem is stronger, its proof is easier: it consists in the explicit construction of a group with the required properties.

In **2** we discuss the problem in a more general context. It arises out of questions on the existence of the 'generalized free product with amalgamated

* Received April 1, 1949.

[1] An element of order zero is understood to generate an infinite cycle.

subgroups dealt with in an earlier paper (cf. [4][2]), and closely related to this, R. Baer's investigations on the imbeddability of a certain type of incomplete group called 'amalgam' in a group (cf. [1]). In the terminology of [4] our result proves the existence of the generalized free product of cyclical groups, thus filling a gap left at that time (cf. [4], **10**, p. 624).

In **3** the general question is reduced to the simpler case of a finite number of generators a_i. These may then be assumed to be all of finite, or all of infinite order. The next two paragraphs deal separately with these cases: The compatibility conditions are given in arithmetical form; this allows one to deal with the case of cycles of finite order; the case of infinite cycles is then reduced to the former.

Finally, in **6**, the result is extended to prove the existence of the generalized free sum of *locally* cyclic groups.

I wish to thank Professor Baer and my husband for many helpful remarks, and in particular the referee for a suggestion which led to a substantial simplification of the proof.

2. The context. We recall briefly the definition of a generalized free product with amalgamated subgroups.

Groups $\mathfrak{G}_\alpha$ are given (finite or infinite in number), with subgroups $\mathfrak{U}_{\alpha\beta} \subseteq \mathfrak{G}_\alpha$ such that $\mathfrak{U}_{\alpha\beta}$ and $\mathfrak{U}_{\beta\alpha}$ are isomorphic. Denote by $I_{\alpha\beta}$ a fixed isomorphism mapping $\mathfrak{U}_{\alpha\beta}$ on $\mathfrak{U}_{\beta\alpha}$, and by $I_{\beta\alpha}$ its inverse. If $u_{\alpha\beta}$ is any element of $\mathfrak{U}_{\alpha\beta}$, denote the corresponding element of $\mathfrak{U}_{\beta\alpha}$ by $u_{\beta\alpha}$,

$$u_{\beta\alpha} = I_{\alpha\beta}u_{\alpha\beta} \quad \text{and} \quad u_{\alpha\beta} = I_{\beta\alpha}u_{\beta\alpha}.$$

Introduce in the free product of the groups $\mathfrak{G}_\alpha$ all relations $u_{\alpha\beta} = u_{\beta\alpha}$. If in the resulting group the groups $\mathfrak{G}_\alpha$ remain 'intact,' i. e. isomorphically represented, such that moreover the intersection of $\mathfrak{G}_\alpha$ and $\mathfrak{G}_\beta$ is exactly $\mathfrak{U}_{\alpha\beta} = \mathfrak{U}_{\beta\alpha}$, then this group is called the generalized free product $\mathfrak{G}$ of the groups $\mathfrak{G}_\alpha$ with amalgamated subgroups $\mathfrak{U}_{\alpha\beta}$.

Clearly, for this construction to be successful, the subgroups $\mathfrak{U}_{\alpha\beta}$ and the isomorphisms $I_{\alpha\beta}$ must fit together; more precisely:

2.1. *For the existence of the generalized free product it is necessary that the three groups*

$$\mathfrak{U}_{\alpha\beta} \cap \mathfrak{U}_{\alpha\gamma} = \mathfrak{U}_{\alpha\beta\gamma} \text{ in } \mathfrak{G}_\alpha,$$
$$\mathfrak{U}_{\beta\gamma} \cap \mathfrak{U}_{\beta\alpha} = \mathfrak{U}_{\beta\gamma\alpha} \text{ in } \mathfrak{G}_\beta,$$
$$\mathfrak{U}_{\gamma\alpha} \cap \mathfrak{U}_{\gamma\beta} = \mathfrak{U}_{\gamma\alpha\beta} \text{ in } \mathfrak{G}_\gamma$$

are isomorphic, such that, moreover,

[2] Numbers in square brackets refer to the list of references at the end of the paper.

2.2. $I_{\beta\gamma}(I_{\alpha\beta}u_{\alpha\beta\gamma}) = I_{\alpha\gamma}u_{\alpha\beta\gamma}$ *for every element* $u_{\alpha\beta\gamma}$ *of* $\mathfrak{U}_{\alpha\beta\gamma}$.

These conditions are in general not sufficient for the existence of $\mathfrak{G}$ (cf. [4], 3. 0 and 3. 2, pp. 595, 596).

We return to the group defined in **1**. Property 1. 3 is equivalent with the existence of the generalized free product of the cyclical groups (a_i) of order m_i with amalgamated subgroups $(a_i^{e_{ij}})$. The necessary conditions 2. 1 and 2. 2 obviously become arithmetical conditions for the numbers m_i and e_{ij}—the 'compatibility conditions' mentioned in the introduction—which in this case will be shown to be also sufficient for the existence of the generalized free product. The explicit arithmetical form of these conditions does not concern us until later (**4**, **5**). Instead we consider the conditions 2. 1 and 2. 2 from a different angle.

A system S of elements is called an incomplete group if a multiplication is defined such that any ordered pair of elements has at most one product, right-division and left-division each have at most one solution, and as far as multiplication is defined it is associative.

Let S consist in particular of all the elements of all the groups $\mathfrak{G}_\alpha$, with amalgamations $u_{\alpha\beta} = u_{\beta\alpha}$ carried out, and with multiplication defined as in the individual groups for any two elements which belong to one and the same group $\mathfrak{G}_\alpha$, and undefined otherwise. This system S is an incomplete group if, and only if, the conditions 2. 1 and 2. 2 are satisfied (cf. [4], **4**, p. 597).

This type of incomplete group we call, with R. Baer, an 'amalgam' (cf. [1], **6**, p. 728). The groups $\mathfrak{G}_\alpha$ we call its 'constituents.'

Remark. One and the same amalgam may be representable as an amalgam of groups in different ways, even involving a finite number of groups one way, but infinitely many the other way. We are, however, interested not in the amalgam as abstract algebraic system, but in given groups formed into an amalgam by virtue of given amalgamations.

The generalized free product of the groups $\mathfrak{G}_\alpha$ may now be described as the group 'freely generated' by the amalgam of these groups. It exists if, and only if, the amalgam is imbeddable in a group (cf. [4], 4. 0, p. 598).

As mentioned already, not every amalgam is imbeddable. R. Baer even gives an example of an amalgam of four abelian groups which is not imbeddable (cf. [1], p. 728). Our result will show that every amalgam of cyclical groups is imbeddable.

Grace E. Bates has generalized the above concepts as follows (cf. [2], Appendix, for more precise definitions): Let all the groups in the amalgam

have property P. The generalized P-free product—if it exists—is the group with property P which contains and is generated by the amalgam, and is such that every other group with the same properties is a homomorphic image of it. If, in particular, property P is taken as 'being abelian,' we are led to the generalized free sum $\mathfrak{A}$ of the abelian groups $\mathfrak{A}_\alpha$. Again, $\mathfrak{A}$ exists if, and only if, the amalgam A of the groups $\mathfrak{A}_\alpha$ is imbeddable in an abelian group.

Clearly, the existence of the generalized free sum implies the existence of the generalized free product. Whether the converse is true in general, is an unsolved problem.[2a] In our special case however it is true: As we are going to imbed any given amalgam of cyclical groups in an *abelian* group, we prove the existence of the generalized free sum, as well as that of the generalized free product, of cyclical groups.

3. Preliminaries. In an amalgam of given groups any subsystem of these groups forms a subamalgam. R. Baer shows that:

An amalgam of an infinite number of groups is imbeddable if every subamalgam formed by a finite number of its constituents is imbeddable. In other words:

Imbeddability of an amalgam is a 'property of finite character.'

It follows similarly:

3. 0. *Imbeddability of an amalgam of abelian groups in an abelian group is a property of finite character.*

The proof is immediate and is omitted.[3]

DEFINITION. *The amalgam A with constituent groups $\mathfrak{G}_\alpha$ is called reducible if there exist in A two proper subamalgams A_1 and A_2 with the following properties:*

(i) *neither A_1 nor A_2 consists of the unit element only;*

(ii) *every constituent $\mathfrak{G}_\alpha$ of A belongs to either A_1 or A_2, but not to both;*

(iii) *every $\mathfrak{G}_\alpha$ in A_1 has unit meet with every $\mathfrak{G}_\beta$ in A_2.*

A is called irreducible if no such subamalgams exist in A. An amalgam consisting of one group only is trivially irreducible.

[2a] [Added in proof, September 1950.] This question has now been settled: I have constructed an amalgam of five abelian groups, imbeddable in a group, but not imbeddable in an *abelian* group. (To appear elsewhere.)

[3] Let Σ be a finite set of identical relations. Then one proves equally easily: Imbeddability of an amalgam of groups whose elements satisfy the relations of Σ in another such group is a property of finite character.

Remark. Reducibility of an amalgam A is actually seen to be independent of the particular way in which A is represented as amalgam of constituent groups. We shall, however, not make use of this fact.

3. 11. *Let A be a reducible amalgam of a finite number of abelian constituents, and let A_1 and A_2 be subamalgams satisfying the conditions* (i)-(iii) *above. If A_1 and A_2 are imbeddable in abelian groups $\mathfrak{A}_1$ and $\mathfrak{A}_2$ respectively, then A is imbeddable in an abelian group, viz. the direct sum $\mathfrak{A}$ of $\mathfrak{A}_1$ and $\mathfrak{A}_2$.*

The proof is obvious. It now follows immediately:

3. 12. *A reducible amalgam of a finite number of abelian constituents is imbeddable in an abelian group if this is true for every irreducible subamalgam.*

Applied to our original problem, 3. 0 and 3. 12 show that we may restrict ourselves to irreducible amalgams formed by a finite number of cycles. But then clearly, either all cycles in the amalgam are of finite order, or they are all of infinite order.

3. 2. *Let A be an amalgam of a finite number of cycles which are either all infinite, or all finite and their orders powers of one and the same prime. If A is irreducible, any two cycles in A have non-trivial meet.*

Proof. Under the assumptions of 3. 2 every constituent group of A has the property that any two non-trivial subgroups have non-trivial meet. Hence 3. 2 follows from the definition of irreducibility in conjunction with 2. 1.

Remark. It follows from a result by R. Baer (cf. [1], **9**, p. 742) that an amalgam of cycles whose orders are all powers of one and the same prime is imbeddable. The case of an amalgam of cycles of arbitrary finite orders can be reduced to this case by splitting up each cycle into a direct sum of cycles of prime power orders. As we wish to prove imbeddability in an abelian group, and besides are interested also in the conditions for given amalgamations 1. 2 to define an amalgam, we shall not rely on Baer's result.

To provide a basis for induction, we note:

3. 31. *Every amalgam of two abelian constituents is imbeddable in an abelian group.*

In fact, more can easily be shown; in obvious terminology (cf. [3], **2**):

3. 32. *The direct product of two groups with one amalgamated subgroup exists provided only the amalgamated subgroup belongs to the centre of each group.*

We are now ready to prove the existence of an abelian group containing a given irreducible amalgam of a finite number of cycles. As we shall now be dealing with abelian groups only, we shall write the groups additively.

DEFINITION. *The system of numbers* e_{ij} $(i, j = 1, \cdots, n)$ *is called compatible with elements of order* m_i $(i = 1, \cdots, n)$, *if the relations* $e_{ij}a_i = e_{ji}a_j$ *and* $m_i a_i = 0$ *for all* i *and* $j \neq i$ *define an amalgam of the* n *cycles* (a_i) *of orders* m_i.

3. 4. *The system of numbers* e_{ij} *is compatible with elements of order* m_i *if, and only if, the subgroups and amalgamations defined by it satisfy* 2. 1 *and* 2. 2.

The general pattern of the subsequent proofs is as follows: From the compatibility of the system e_{ij} with elements of order m_i we derive arithmetical conditions satisfied by the numbers e_{ij} and m_i. These *necessary* conditions for compatibility are then shown to imply the existence of an abelian group containing the given amalgam. In this way the arithmetical conditions are shown to be also sufficient, and simultaneously the imbeddability of the amalgam is established.

4. Amalgams of finite cycles. We deal with this case in three steps: first we assume the cycles to be of equal prime power order p^μ. Then the case of n cycles of arbitrary prime power orders p^{μ_i} $(i = 1, \cdots, n)$ is reduced to the previous case; finally the case of n cycles of arbitrary finite orders m_i is reduced to the case of cycles of prime power orders.

Throughout the rest of the paper we write $r \vee s$ for the positive greatest common divisor of two integers r and s. Also, if p is any prime, $p^{\lambda(r)}$ the highest power of p dividing the integer r, we call $\lambda(r)$ the p-contents of r, i. e. $r = p^{\lambda(r)} r_1$, where $r_1 \vee p = 1$. (No ambiguity will arise from the fact that this notation $\lambda(r)$ does not indicate the particular prime p under consideration).

Given are n elements a_i with $p^\mu a_i = 0$ and $e_{ij}a_i = e_{ji}a_j$. We put

4. 01. $\qquad \lambda(e_{ij}) = \epsilon_{ij}$, i. e. $e_{ij} = d_{ij}p^{\epsilon_{ij}}$ with $d_{ij} \vee p = 1$.

Since only irreducible amalgams need be considered, we shall assume throughout (because of 3. 2):

4. 02. $\qquad \epsilon_{ij} < \mu$ for all pairs i, j.

4. 1. LEMMA. *If the system of numbers* $e_{ij} = d_{ij}p^{\epsilon_{ij}}$ *is compatible with* n *elements of order* p^μ, *then*

4.10. $\epsilon_{ij} = \epsilon_{ji}$ for all pairs i, j;

4.11. *of any three exponents* ϵ_{ij}, ϵ_{jk}, ϵ_{ki} *two are equal, the third not greater than these two*;

4.12. *for any three suffixes* i, j, k

$$d_{ij}d_{jk}d_{ki} \equiv d_{ji}d_{kj}d_{ik} \bmod p^{\mu-\epsilon} \text{ where } \epsilon = \operatorname{Max}(\epsilon_{ij}, \epsilon_{jk}, \epsilon_{ki}).$$

Proof. Since the relations $e_{ij}a_i = e_{ji}a_j$ define an amalgam of the cycles (a_i) of order p^μ, the amalgamated subgroups are isomorphic, which implies 4.10.

To prove 4.11, we use the fact that the amalgamations satisfy 2.1. If for a fixed triplet i, j, k the three exponents ϵ_{ij}, ϵ_{jk}, and ϵ_{ki} are not equal, assume without loss of generality that $\epsilon_{ij} > \epsilon_{ik}$. Thus the amalgamated subgroup $(e_{ij}a_i)$ of (a_i) is contained in $(e_{ik}a_i)$, and their intersection, i. e. the intersection of (a_i), (a_j), and (a_k), is $(e_{ij}a_i)$ of order $p^{\mu-\epsilon_{ij}}$. But this same intersection, formed in (a_k), is the intersection of $(e_{ki}a_k)$ and $(e_{kj}a_k)$. For this to have the same order $p^{\mu-\epsilon_{ij}}$, necessarily $\epsilon_{kj} = \epsilon_{ij}$, because of 4.10. This proves 4.11.

To prove 4.12, we note that the amalgamations satisfy 2.2, i. e. are consistent with each other. Assume $\epsilon_{ij} \geq \epsilon_{ik}$, and consider the element $p^{\epsilon_{ij}}d_{ij}d_{jk}d_{ki}a_i$ of (a_i). Repeated application of the given relations (noting $p^{\epsilon_{ij}}d_{ij} = e_{ij}$, etc.) gives: $p^{\epsilon_{ij}}d_{ij}d_{jk}d_{ki}a_i = p^{\epsilon_{ij}}d_{ji}d_{jk}d_{ki}a_j$, which is $p^{\epsilon_{ij}}d_{ji}d_{kj}d_{ki}a_k$ (as $\epsilon_{ij} = \epsilon_{jk}$), or $p^{\epsilon_{ij}}d_{ji}d_{kj}d_{ik}a_i$ (as $\epsilon_{ij} \geq \epsilon_{ik}$); hence these two multiples of a_i represent the same element of (a_i), and 4.12 follows.

4.2. LEMMA. *If the system of numbers* e_{ij} *and* p^μ *satisfies the conditions* 4.10-4.12, *then there exists an abelian group* $\mathfrak{A}$ *generated by elements* $a_1, \cdots, a_n$ *of orders* p^μ *such that in* $\mathfrak{A}$ *the intersection of* (a_i) *and* (a_j) *is generated by the element* $e_{ij}a_i = e_{ji}a_j$ *for all* i *and* j.

Proof. If $n = 2$, the lemma is true by 3.3. We shall, therefore, prove it by induction with respect to n.

Assume that $\mathfrak{A}_1$ is an abelian group generated by elements $a_1, \cdots, a_{n-1}$ of orders p^μ, such that in $\mathfrak{A}_1$

4.20. $(a_i) \cap (a_j) = (e_{ij}a_i = e_{ji}a_j)$ for $i, j = 1, \cdots, n-1$.

We construct an abelian group $\mathfrak{A}$ containing the given amalgam of n cycles as free sum of $\mathfrak{A}_1$ and a cycle (a_n) of order p^μ with a suitable amalgamation.

Amongst the intersections (a_n) is to have with the cycles (a_i) $(i \leq n-1)$, there is a maximal one which may be assumed to be its intersection with (a_1), i. e. the subgroup of order $p^{\mu-\epsilon_{n1}}$ generated by $e_{n1}a_n$. The numbers e_{ij} satisfy the conditions 4. 10-4. 12; because of the first of these, the subgroup $(e_{1n}a_1)$ of (a_1), i. e. of $\mathfrak{A}_1$, is isomorphic with $(e_{n1}a_n)$. Therefore the free sum $\mathfrak{A}$ of $\mathfrak{A}_1$ and (a_n) with the amalgamation $e_{1n}a_1 = e_{n1}a_n$ exists. We prove that $\mathfrak{A}$ has the desired properties.

Since $\mathfrak{A}_1$ is isomorphically contained in $\mathfrak{A}$, the intersection of any two of the first $n-1$ cycles (a_i) is given by 4. 20 also in $\mathfrak{A}$. All we have to prove therefore is that also the intersection of (a_i) and (a_n) is the group generated by $e_{in}a_i = e_{ni}a_n$, for $i = 1, \cdots, n-1$. For $i = 1$ this follows from the construction of $\mathfrak{A}$, for $i > 1$ we deduce it from 4. 1.

First we show that $(a_i) \cap (a_n)$ is the subgroup of order $p^{\mu-\epsilon_{in}}$ of (a_n), $(2 \leq i \leq n-1)$.

Now $\mathfrak{A}_1 \cap (a_n) = (e_{1n}a_1 = e_{n1}a_n)$, and $(a_i) \subseteq \mathfrak{A}_1$. Hence

$$(a_i) \cap (a_n) = (a_i) \cap (e_{1n}a_1) \cap (a_n) = (a_i) \cap (e_{1n}a_1),$$

so that

4. 21. $\quad (a_i) \cap (a_n) = (a_i) \cap (a_1)$, if $(a_i) \cap (a_1) \subset (e_{1n}a_1)$,

4. 22. $\quad (a_i) \cap (a_n) = (e_{1n}a_1)$, if $(a_i) \cap (a_1) \supseteq (e_{1n}a_1)$.

But $(a_1) \cap (a_1)$ is of order $p^{\mu-\epsilon_{i1}}$, $(e_{1n}a_1)$ is of order $p^{\mu-\epsilon_{1n}}$. Hence, in the first case (4. 21), $\epsilon_{i1} > \epsilon_{1n}$, so that by 4. 11 $\epsilon_{in} = \epsilon_{i1}$; in the second case (4. 22), $\epsilon_{i1} \leq \epsilon_{1n}$, hence, by 4. 11, $\epsilon_{in} = \epsilon_{1n}$. Thus in both cases $(a_i) \cap (a_n)$ has the desired order.

It remains to prove that the correct multiples $e_{in}a_i = d_{in}p^{\epsilon_{in}}a_i$ and $e_{ni}a_n = d_{ni}p^{\epsilon_{in}}a_n$ of $p^{\epsilon_{in}}a_i$ and $p^{\epsilon_{in}}a_n$ respectively are equal in $\mathfrak{A}$.

In the following lines, d_{ij}^{-1} denotes the unique inverse mod p^{μ} of d_{ij}.

In the first case (4. 21), $(a_i) \cap (a_n)$ is the subgroup $(p^{\epsilon_{in}}a_i)$ of (a_i), which, because of $\epsilon_{in} = \epsilon_{i1}$, is also generated by

$$e_{in}a_i = d_{in}p^{\epsilon_{in}}a_i = d_{in}d_{i1}^{-1}d_{i1}p^{\epsilon_{in}}a_i = d_{in}d_{i1}^{-1}d_{1i}p^{\epsilon_{in}}a_1.$$

Also, as $\epsilon_{in} > \epsilon_{1n}$,

$$e_{ni}a_n = d_{ni}p^{\epsilon_{in}}a_n = d_{ni}d_{n1}^{-1}d_{n1}p^{\epsilon_{in}}a_n = d_{ni}d_{n1}^{-1}d_{1n}p^{\epsilon_{in}}a_1;$$

therefore by 4. 12: $e_{in}a_i = e_{ni}a_n$.

In the second case (4. 22) the proof is similar and will be omitted.

Next we turn to the case of n elements a_i of orders p^{μ_i}, and $e_{ij}a_i = e_{ji}a_j$. The p-contents of e_{ij} is denoted by ϵ_{ij} as before.

4. 3. LEMMA. *If the system of numbers $e_{ij} = d_{ij}p^{\epsilon_{ij}}$ is compatible with n elements of orders p^{μ_i}, and if*

4. 30. $$\mu = \operatorname{Max} \mu_i \text{ and } \delta_{ij} = \epsilon_{ij} + \mu - \mu_i,$$

then

4. 31. $$\delta_{ij} = \delta_{ji} \text{ for all pairs } i, j;$$

4. 32. *of any three numbers δ_{ij}, δ_{jk}, δ_{ki} two are equal, the third not greater than these two;*

4. 33. *for any three suffixes i, j, k:*

$$d_{ij}d_{jk}d_{ki} \equiv d_{ji}d_{kj}d_{ik} \bmod p^{\mu-\delta} \text{ where } \delta = \operatorname{Max}(\delta_{ij}, \delta_{jk}, \delta_{ki}).$$

4. 4. LEMMA. *If the system of numbers e_{ij} and p^{μ_i} satisfies the conditions* 4. 30-4. 33, *then there exists an abelian group $\mathfrak{A}$ generated by elements a_i of orders p^{μ_i} $(i = 1, \cdots, n)$, such that in $\mathfrak{A}$ the intersection $(a_i) \bigcap (a_j)$ is generated by the element $e_{ij}a_i = e_{ji}a_j$, for all i and j.*

Proof of 4. 3. Consider n elements $b_1, \cdots, b_n$ of equal order p^{μ}, and in each cycle (b_i) the element $a_i = p^{\mu-\mu_i}b_i$ $(i = 1, \cdots, n)$. Since a_i is of order p^{μ_i}, the relations $e_{ij}a_i = e_{ji}a_j$ define an amalgam of the $a_i = p^{\mu-\mu_i}b_i$, i. e. the relations $e_{ij}p^{\mu-\mu_i} \cdot b_i = e_{ji}p^{\mu-\mu_j} \cdot b_j$ define an amalgam of the cycles b_i. Thus the system of numbers $e_{ij}p^{\mu-\mu_i}$ is compatible with n elements of equal order p^{μ}, and 4. 3 follows from 4. 1.

Proof of 4. 4. With the same notation as in the proof of 4. 3, we have: under the assumption of 4. 4, the system of numbers $p^{\mu-\mu_i}e_{ij}$ and p^{μ} satisfies the conditions which, by 4. 2, are sufficient for the existence of an abelian group $\mathfrak{B}$ generated by n elements b_i of equal order p^{μ} such that $(b_i) \bigcap (b_j)$ is generated by $e_{ij}p^{\mu-\mu_i}b_i = e_{ji}p^{\mu-\mu_j}b_j$. But then clearly the elements $a_i = p^{\mu-\mu_i}b_i$ of $\mathfrak{B}$ and the subgroup $\mathfrak{A}$ of $\mathfrak{B}$ generated by them satisfy the requirements of the lemma.

Finally, let $a_1, \cdots, a_n$ be elements of arbitrary orders $m_1, \cdots, m_n$, and let again $e_{ij}a_i = e_{ji}a_j$. For any prime p let $\lambda_p(m_i)$ denote the p-contents of m_i.

4. 5. LEMMA. *The system of numbers e_{ij} is compatible with n elements of orders m_i if, and only if, it is compatible with n elements of orders $p^{\lambda_p(m_i)}$ for every prime p.*

4. 6. THEOREM. *Every irreducible amalgam of a finite number of cycles of finite order is imbeddable in an abelian group.*

Proof of 4. 5 *and* 4. 6. Clearly 4. 5 is void for any p which does not divide any of the numbers m_i. Let $p_1, \cdots, p_r$ be all the primes dividing at least one of the m_i. The cycle (a_i) of order m_i is the direct sum of r cycles $(b_{i\rho})$ of order $m_{i\rho}$, where $m_{i\rho} = p_\rho^{\lambda_\rho(m_i)}$, and $\rho = 1, \cdots, r$:

$$a_i = b_{i1} + b_{i2} + \cdots + b_{ir} \text{ with } m_{i\rho} b_{i\rho} = 0,$$

for $i = 1, \cdots, n$. The amalgamations $e_{ij}a_i = e_{ji}a_j$ induce the same amalgamations $e_{ij}b_{i\rho} = e_{ji}b_{j\rho}$ for each fixed value of ρ. If the cycles (a_i) form an amalgam A by means of the given amalgamations, then clearly the cycles $(b_{i\rho})$ (for $i = 1, \cdots, n$) form an amalgam B_ρ by means of the same amalgamations, and conversely. Also, the amalgams B_ρ may be imbedded in abelian groups $\mathfrak{B}_\rho$, by 4. 3 and 4. 4. But then A is imbeddable in the direct sum $\mathfrak{A}$ of the abelian groups $\mathfrak{B}_\rho$.

5. Amalgams of infinite cycles. Again we begin by deriving necessary conditions for the numbers e_{ij} to be compatible with n elements of infinite order.

Given are n elements of infinite order, $a_1, \cdots, a_n$ with $e_{ij}a_i = e_{ji}a_j$. Only irreducible amalgams need be considered, hence we shall assume throughout this paragraph (because of 3. 2).

5. 0. $$e_{ij} \neq 0 \text{ for all } i, j.$$

If the given relations define an amalgam of the n cycles, the subgroups and amalgamations defined by them satisfy 2. 1 and 2. 2. The former gives nothing new in this case; from the latter we deduce

5. 1. LEMMA. *If the system of numbers e_{ij} is compatible with n elements of infinite order, then for any three suffixes i, j, k:*

5. 10. $$|\, e_{ij} \,|\, (e_{kj} \vee e_{ki}) = |\, e_{kj} \,|\, (e_{ij} \vee e_{ik}) ;$$

5. 11. $$e_{ij}e_{jk}e_{ki} = e_{ji}e_{kj}e_{ik}.$$

Proof. 5. 10, written in the form $|\, e_{ij} \,|/(e_{ij} \vee e_{ik}) = |\, e_{kj} \,|/(e_{kj} \vee e_{ki})$, simply expresses the fact that the index of the three-suffix meet $(a_i) \cap (a_j) \cap (a_k)$ under the intersection $(a_i) \cap (a_k)$ may be calculated in (a_i) or in (a_k) with the same result. Thus 5. 10 holds in an amalgam.

5. 11 follows again by repeated application of the given amalgamations to the element $e_{ij}e_{jk}e_{ki}a_i$:

$$e_{ij}e_{jk}e_{ki}a_i = e_{ji}e_{jk}e_{ki}a_j = e_{ji}e_{kj}e_{ki}a_k = e_{ji}e_{kj}e_{ik}a_i ;$$

thus these two multiples of a_i are the same element of the amalgam, and therefore identical, as a_i is of infinite order.

The following form of the compatibility condition derived from 5.1 will allow us to reduce the present problem to that of imbedding an amalgam of finite cycles.

We denote the greatest common divisor of e_{ij} and e_{ji} by r_{ij}, so that for all pairs i, j

5.20. $$e_{ij} = c_{ij} r_{ij}, \qquad e_{ji} = c_{ji} r_{ij}, \qquad r_{ij} = r_{ji}$$

and

5.21. $$c_{ij} \vee c_{ji} = 1.$$

5.3. LEMMA. *If the system of numbers e_{ij} is compatible with n elements of infinite order, then*

5.30 *there exist n non-zero integers $x_1, \cdots, x_n$ with greatest common divisor unity, such that $c_{ij} = x_j/(x_i \vee x_j)$ for all pairs i, j;*

5.31 *the numbers e_{ij} are also compatible with n elements of finite orders $m_i = r h_i$, where r is the least common multiple of all numbers r_{ij}, and h_i is the least common multiple of the $n-1$ numbers c_{ij} $(j = 1, \cdots, i-1, i+1, \cdots, n)$.*

Proof of 5.30. We consider the $n(n-1)/2$ linear equations

$$e_{ij} x_i - e_{ji} x_j = 0 \qquad (i, j = 1, \cdots, n; i \neq j).$$

The matrix of coefficients of this system we denote by E. E has $n(n-1)/2$ rows and n columns; the order in which the rows are taken, is of course irrelevant. Each row contains exactly two non-zero elements, e_{ij} and $-e_{ji}$; their column is given by their first suffix. We prove:

5.4. *The matrix E has rank $n-1$.*

Proof. By means of 5.11 taken for the suffixes $1, i, j$, the equation (i, j)—in obvious notation—follows from the two equations $(1, i)$ and $(1, j)$. Hence all equations follow from the $n-1$ equations $e_{1i} x_1 - e_{i1} x_i = 0$ $(i = 2, \cdots, n)$, and these are clearly independent.

Thus the equations have essentially one non-trivial solution, and necessarily all $x_i \neq 0$ by 5.0. Now let $x_1, \cdots, x_n$ denote the solution consisting of integers without common divisor (unique but for a common change of

sign). Since with $e_{ij}x_i - e_{ji}x_j = 0$ also $c_{ij}x_i - c_{ji}x_j = 0$, these numbers x_i satisfy the requirements of 5. 30 because of 5. 21.

Proof of 5. 31. We bring this proof in three steps.

1. By 4. 5, it is sufficient to prove that the numbers e_{ij} are compatible with n elements of orders $p^{\lambda(m_i)}$ for any prime p. From now on let p be an arbitrary fixed prime. Put

$$\lambda(e_{ij}) = \epsilon_{ij}, \qquad \lambda(m_i) = \mu_i, \qquad \text{Max}\, \mu_i = \mu.$$

Applying Lemmas 4. 3 and 4. 4 we see that we need only prove that the conditions 4. 31 and 4. 32 are satisfied by the numbers $\delta_{ij} = \epsilon_{ij} + \mu - \mu_i$. The third condition, 4. 33, is then trivially satisfied since we even have $e_{ij}e_{jk}e_{ki} = e_{ji}e_{kj}e_{ik}$.

2. We first calculate the numbers ϵ_{ij} and δ_{ij}. Let

5. 50. $\qquad \lambda(x_i) = \alpha_i \qquad (i = 1, \cdots, n; x_i \text{ defined by } 5.30).$

We assume without loss of generality

5. 51. $$\alpha_1 = 0 \leq \alpha_2 \leq \cdots \leq \alpha_n.$$

5. 52. Hence $\lambda(c_{ij}) = \alpha_j - \alpha_i$, if $i < j$; $\lambda(c_{ij}) = 0$, if $i > j$.

Put $\lambda(r_{ij}) = \rho_{ij}$ and $\text{Max}\, \rho_{ij} = \rho$.

Then we obtain for the p-contents of the numbers e_{ij}:

5. 53. $\qquad \epsilon_{ij} = \alpha_j - \alpha_i + \rho_{ij}$, if $i < j$; $\epsilon_{ij} = \rho_{ij}$, if $i > j$;

thus for $\mu_i = \lambda(m_i)$ from the definition of m_i, and $\mu = \text{Max}\, \mu_i$:

5. 54. $\qquad \mu_i = \text{Max}_j(\alpha_j - \alpha_i) + \rho = \alpha_n - \alpha_i + \rho,$

and $\mu = \text{Max}_i\, \mu_i = \alpha_n - \alpha_1 + \rho = \alpha_n + \rho$.

Hence finally:

5. 55. $\qquad \delta_{ij} = \alpha_j + \rho_{ij}$, if $i < j$, and $\delta_{ij} = \alpha_i + \rho_{ij}$, if $i > j$.

Thus $\delta_{ij} = \delta_{ji}$ follows immediately.

3. To prove 4. 32 we need information on the numbers ρ_{ij}. We assume without loss of generality that $i < j < k$. Then

5. 56. $\qquad \delta_{ij} = \alpha_j + \rho_{ij}, \qquad \delta_{jk} = \alpha_k + \rho_{jk}, \qquad \delta_{ki} = \alpha_k + \rho_{ki}.$

To find the relation between the numbers ρ_{ij}, ρ_{jk}, ρ_{ki}, we use the relations 5. 10. for the triplets (i, j, k) and (j, i, k) of suffixes. By their nature they imply the corresponding conditions for the p-contents of the numbers e_{ij}, viz.:

$$\epsilon_{ij} + \operatorname{Min}(\epsilon_{kj}, \epsilon_{ki}) = \epsilon_{kj} + \operatorname{Min}(\epsilon_{ij}, \epsilon_{ik})$$

and

$$\epsilon_{ji} + \operatorname{Min}(\epsilon_{kj}, \epsilon_{ki}) = \epsilon_{ki} + \operatorname{Min}(\epsilon_{ji}, \epsilon_{jk}).$$

To save writing, we put for the moment $\rho_{ij} = \beta$, $\rho_{jk} = \gamma$, $\rho_{ki} = \delta$.

Then we have, because of 5. 53 and $i < j < k$:

5. 57. $$\alpha_j - \alpha_i + \beta + \operatorname{Min}(\gamma, \delta) = \gamma + \operatorname{Min}(\alpha_j - \alpha_i + \beta, \alpha_k - \alpha_i + \delta)$$

and

$$\beta + \operatorname{Min}(\gamma, \delta) = \delta + \operatorname{Min}(\beta, \alpha_k - \alpha_j + \gamma);$$

hence, subtracting $\alpha_j - \alpha_i$ from both sides of the first relation,

5. 58. $$\beta + \operatorname{Min}(\gamma, \delta) = \gamma + \operatorname{Min}(\beta, \alpha_k - \alpha_j + \delta) = \delta + \operatorname{Min}(\beta, \alpha_k - \alpha_j + \gamma).$$

We show

5. 59. If $\gamma = \delta$, then $\beta \leq \alpha_k - \alpha_j + \delta$;
if $\gamma < \delta$, then $\beta = \alpha_k - \alpha_j + \delta$;
if $\gamma > \delta$, then $\beta = \alpha_k - \alpha_j + \gamma$.

Proof. The first possibility is obvious. Otherwise, let $\gamma < \delta$; then the first of the three expressions 5. 58 equals $\beta + \gamma$; for the last to have this value, $\alpha_k - \alpha_j + \gamma \leq \beta$ is necessary and then $\alpha_k - \alpha_j + \gamma + \delta = \beta + \gamma$ follows, and satisfies 5. 58. The last possibility follows by interchange of γ and δ: 5. 58 is symmetrical in γ and δ.

But the relations 5. 59 for the numbers ρ_{ij}, ρ_{jk}, ρ_{ki} together with 5. 56 and $\alpha_i \leq \alpha_j \leq \alpha_k$ prove 4. 32 for δ_{ij}, δ_{jk}, and δ_{ki}.

Now we obtain easily:

5. 6. THEOREM. *An irreducible amalgam of a finite number of infinite cycles is imbeddable in an abelian group.*

Proof. If the relations $e_{ij}a_i = e_{ji}a_j$ define an amalgam, the numbers e_{ij} satisfy the conditions of 5. 3. By 5. 32 and 4. 5, there exists an abelian group $\mathfrak{B}$ generated by n elements b_i, such that in $\mathfrak{B}$ the intersection of (b_i) and (b_j) is generated by the element $e_{ij}b_i = e_{ji}b_j$. Since $r_{ij} = e_{ij} \vee e_{ji}$ divides the orders of both b_i and b_j, the element $e_{ij}b_i - e_{ji}b_j$ of $\mathfrak{B}$ has the exact order r_{ij}.

We now form the direct sum $\mathfrak{A}$ of $\mathfrak{B}$ with one infinite cycle (a). Let $x_1, \cdots, x_n$ be integers as defined in 5. 30; the elements $a_i = x_i a + b_i$ of $\mathfrak{A}$ $(i = 1, \cdots, n)$ are of infinite order. Also because of the infinite component $x_i a$ of a_i, the intersection $(a_i) \cap (a_j)$ is generated by a multiple $rc_{ij}a_i = rc_{ji}a_j$

of $c_{ij}a_i$ and $c_{ji}a_j$. Because of the properties of b_i and b_j in $\mathfrak{B}$, $r = r_{ij}$. Thus the intersection of (a_i) and (a_j) is generated by the element $e_{ij}a_i = e_{ji}a_j$. This completes the proof.

We add a remark on the nature of the generalized free sum in the case considered in this paragraph:

5. 7. THEOREM. *The generalized free sum of n infinite cycles, no two of which have zero meet, is the direct sum of one infinite cycle and at most $n-1$ finite cycles. It is cyclical if, and only if, the amalgamations $e_{ij}a_i = e_{ji}a_j$ satisfy $e_{ij} \vee e_{ji} = 1$ for all pairs i, j.*

Proof. Consider again the matrix E defined as the matrix of coefficients of the system of linear equations $e_{ij}x_i - e_{ji}x_j = 0$. This is clearly the characteristic matrix associated with the abelian group generated by $a_1, \cdots, a_n$ with relations $e_{ij}a_i = e_{ji}a_j$, i. e. with the generalized free sum in question. It was shown to have rank $n-1$ (cf. 5. 4) which proves the first part of the theorem.

Also, if $e_{ij} \vee e_{ji} = r_{ij} \neq 1$, for some pair i, j, then the element $c_{ij}a_i - c_{ji}a_j$ is of order $r_{ij} \neq 1$, and the generalized free sum is certainly not an infinite cycle.

To prove that the group is cyclical if $e_{ij} \vee e_{ji} = 1$, i. e. $e_{ij} = c_{ij}$ in the notation used in Lemma 5. 3, it is sufficient to show that the greatest common divisor of all determinants of order $n-1$ which can be formed from E is unity; or again:

If p is any prime, there is at least one determinant of order $n-1$ in E which is not divisible by p.

By 5. 30 the non-zero elements of E are of the form $c_{ij} = x_j/(x_i \vee x_j)$. We put again $\lambda(x_i) = \alpha_i$, and may assume $\alpha_1 = 0 \leq \alpha_2 \leq \cdots \leq \alpha_n$. Now consider the $n-1$ rows of E containing the elements c_{2j} and $-c_{j2}$ ($j = 1, 3, \cdots, n$; the first suffix indicates the column). One of the non-zero determinants which can be formed from these rows equals, but for the sign,

$$c_{21}c_{32}\cdots c_{n2} = \frac{x_1}{x_1 \vee x_2} \cdot \frac{{x_2}^{n-2}}{(x_2 \vee x_3)\cdots(x_2 \vee x_n)}.$$

As p is prime to x_1, and the p-content of $x_2 \vee x_j$ ($j \geq 3$) is the same as that of x_2, this determinant is prime to p.

6. Amalgams of locally cyclic groups. Let A be an amalgam of groups $\mathfrak{G}_\alpha$ and ϕ_α any set of subgroups of $\mathfrak{G}_\alpha$. By virtue of being subgroups of a group, the groups of the set ϕ_α naturally form an amalgam. Also, all the

groups of all the sets ϕ_α form an amalgam by virtue of the given amalgamations between the groups $\mathfrak{G}_\alpha$ in the amalgam A.[4]

If, in particular, we take as ϕ_α the set of all cyclical subgroups of $\mathfrak{G}_\alpha$, we thus obtain the amalgam of all cyclical subgroups of A. We denote it by A^*. It follows from our previous results that A^* is imbeddable in an abelian group.[5] In general, of course, this does not imply imbeddability of A: A and A^* consist of the same elements; also, two elements whose product is defined in A^*, also have a product in A (and then, of course, the same as in A^*) but, in general, the converse is not true.

We consider the extreme case that A and A^* are identical. This is true if, and only if, any two elements whose product is defined in A, also have a product in A^*, i. e. if, and only if, any two elements belonging to the same group $\mathfrak{G}_\alpha$ belong to, and therefore generate, a cyclical subgroup of $\mathfrak{G}_\alpha$. But this just means that every group $\mathfrak{G}_\alpha$ is locally cyclic. Thus we have proved:

6. 0. LEMMA. *An amalgam of groups $\mathfrak{G}_\alpha$ and the amalgam of all its cyclical subgroups are identical if, and only if, all groups $\mathfrak{G}_\alpha$ are locally cyclic.*

6. 1. THEOREM. *An amalgam of locally cyclic groups is imbeddable in an abelian group.*

UNIVERSITY COLLEGE, HULL.

REFERENCES.

[1] R. Baer, "Free sums of groups and their generalizations," *American Journal of Mathematics*, vol. 71 (1949), pp. 706-742.

[2] Grace E. Bates, "Free loops and nets and their generalizations," *American Journal of Mathematics*, vol. 69 (1947), pp. 499-550.

[3] B. H. Neumann and Hanna Neumann, "A remark on generalized free products," *Journal of the London Mathematical Society*, vol. 25 (1950), pp. 202-204.

[4] Hanna Neumann, "Generalized free products with amalgamated subgroups, I," *American Journal of Mathematics*, vol. 70 (1948), pp. 590-625.

[4] This is a special case of the more general fact that an amalgam (suitably defined) of amalgams of groups is an amalgam of groups.

[5] If A itself is a group, this amalgam A^*, and therefore the abelian group freely generated by it, is an invariant of the group. I owe this remark to the referee.

GENERALIZED EUCLIDEAN SPACE IN TERMS OF A QUASI INNER PRODUCT.*

By LEONARD M. BLUMENTHAL.

---

1. Introduction. Necessary and sufficient conditions of various kinds have been obtained in order that an inner product may be defined in a normed linear space, with the norm and inner product related in the customary manner. Such conditions have been expressed as norm or distance relations by Fréchet [4], Jordan and von Neumann [5], and Aronszajn [1], while other writers have formulated the conditions in terms of orthogonality, etc., for spaces of more than two dimensions.

In a normed linear space S, three operations (subjected to eleven conditions) are assumed as primitive; namely, *addition* (on SS to S), *scalar multiplication* (on RS to S), and *norm* (on S to R, the set of real numbers). If an additional condition is placed on these operations (for example, the condition $\| x+y \|^2 + \| x-y \|^2 = 2(\| x \|^2 + \| y \|^2)$ of Jordan and von Neumann), then a fourth operation, *inner product* (on SS to R), may be defined in S, satisfying five relations connecting the inner product with the three primitive operations of the space.

We solve in this paper the inverse problem. An abstract set Σ is considered in which just *one* primitive operation, on $\Sigma\Sigma$ to R (quasi inner product), is asumed to satisfy *three* conditions. It is proved that if Σ is rich enough in elements to satisfy three existence postulates, then three additional properties of the quasi inner product are necessary and sufficient to permit *defining* in Σ addition, scalar multiplication, and norm so that the resulting space is a normed linear space with an inner product connected with the norm in the usual way. The concluding section of the paper is concerned with the congruent imbedding and vectorial application of Σ onto euclidean and Hilbert spaces, yielding, incidentally, a new set of postulates for Hilbert space (in terms of a single operation).

It is emphasized that our postulates and fundamental definitions are all expressed wholly and explicitly in terms of the quasi inner products of those and only those elements of Σ directly concerned. We shall advert to

---

* Received March 30, 1950. Presented to the American Mathematical Society, December 27, 1949.

this feature in the concluding section of the paper when its significance can be more readily made clear.

The first sections of the paper may be considered as giving a foundation of vector algebra in terms of one operation (quasi inner product). From that point of view our work is related to a brief sketch (without proofs) of a foundation for vector algebra given by Menger [6].

2. The postulates. All of the postulates to be imposed are listed in this section, although they will be introduced later only as needed.

POSTULATES FOR A QUASI INNER PRODUCT.

To each pair of elements x, y of an abstract set $\mathfrak{S}$ there is attached a real number (x, y), called a *quasi inner product,* in conformity with the following agreements:

Q_1 (**Symmetry**). *If* $x, y \in \mathfrak{S}$, *then* $(x, y) = (y, x)$.

Q_2 (**Definiteness**). *For each element* x *of* $\mathfrak{S}$, $(x, x) \geqq 0$.

Q_3 (**Identification**). *If* $x, y \in \mathfrak{S}$ *and* $(x, x) = (x, y) = (y, y)$, *then* $x = y$.

(To avoid misunderstanding, it is pointed out that Q_3 is *not* a definition of equality of elements of $\mathfrak{S}$ which, as a set, is already supplied with a criterion for equality of its elements. As its listing indicates, Q_3 is an *assumption* connecting the set-equality of elements x, y with the *three* numbers (x, x), (x, y), (y, y) attached to the two elements as quasi inner products. On the other hand, we shall show later that Q_3 does, in the presence of Q_1, Q_2, S_2, E_1 (see below) actually define an equivalence relation in $\mathfrak{S}$.)

The set $\mathfrak{S}$ may then be referred to as a quasi inner product space.

SCHWARZ POSTULATES.

If $x_1, x_2, \cdots, x_n \in \mathfrak{S}$, denote by $G(x_1, x_2, \cdots, x_n)$ the Gram determinant $|(x_i, x_j)|$, $(i, j = 1, 2, \cdots, n)$.

S_1. *If* $x_1, x_2 \in \mathfrak{S}$, *then* $G(x_1, x_2) \geqq 0$.

S_2. *If* $x_1, x_2, x_3 \in \mathfrak{S}$ *and* $G(x_1, x_2) = 0$, *then* $G(x_1, x_2, x_3) \geqq 0$.

S_3. *If* $x_1, x_2, x_3, x_4 \in \mathfrak{S}$ *and* $G(x_1, x_2, x_3) = 0$, *then* $G(x_1, x_2, x_3, x_4) \geqq 0$.

It is observed that S_2 and S_3 are *conditional* Schwarz inequalities, while S_1 applies to every pair of elements of $\mathfrak{S}$. The existence postulates that follow merely insure that $\mathfrak{S}$ has enough elements for our purpose.

EXISTENCE POSTULATES.

If $x_1, x_2, \cdots, x_n \in \Sigma$, denote by $B(x_1, x_2, \cdots, x_n)$ the symmetric determinant obtained by bordering $G(x_1, x_2, \cdots, x_n)$ with a row and column of 1's with intersection element 0.

E_1. *There exists at least one element θ of Σ such that $(\theta,x) = 0$ for each element x of Σ.*

E_2. *For each element x of Σ and each real number λ there exists at least one element y of Σ such that $G(x, y) = 0$ and $(x, y) = \lambda \cdot (x, x)$.*

E_3. *If $x, z \in \Sigma$ and $G(x, z) \neq 0$, there exists at least one element y of Σ such that $B(x, y, z) = G(x, y, z) = 0$ and $G(x, y) = G(y, z)$.*

If scalar multiplication and addition are defined in a set, we recall that a real function $((x, y))$ defined over the set of all pairs of elements is a (real) *inner product* provided: (a) $((x, y)) = ((y, x))$, (b) $((x, x)) \geqq 0$, (c) $((x, x)) = 0$ if and only if x is the element of the space that is neutral with respect to addition, (d) $((\lambda \cdot x, y)) = \lambda \cdot ((x, y))$, and (e) $((x + y, z)) = ((x, z)) + ((y, z))$.

The reader is referred to [2] for the definition of a normed linear space.

DEFINITION 2.1. *A normed linear space in which an inner product is defined so that for each element x, the norm $\| x \|$ and inner product are connected by the relation $\| x \| = ((x, x))^{\frac{1}{2}}$, is called a generalized euclidean space.*

3. Development of the system $\{\Sigma: Q_1, Q_2, Q_3, S_2, E_1\}$. We assume throughout this section that Σ is subjected to just the five postulates listed in the heading.

THEOREM 3.1. *There is exactly one element θ of Σ with $(\theta, x) = 0$ for each element x of Σ.*

Proof. If θ, θ' are two such elements then $(\theta, \theta) = (\theta, \theta') = (\theta', \theta') = 0$, and so $\theta = \theta'$ by Q_3. The theorem now follows from E_1.

THEOREM 3.2. *The quasi inner product (x, x) vanishes if and only if $x = \theta$.*

Proof. By E_1, $(\theta, \theta) = 0$, and if $(x, x) = 0$ then $(x, x) = (\theta, x) = (\theta, \theta)$, which implies $x = \theta$ by Q_3.

THEOREM 3.3. *Corresponding to each element x of Σ $(x \neq \theta)$, and each real number λ, there is at most one element y of Σ such that $(x, y) = \lambda \cdot (x, x)$ and $G(x, y) = 0$.*

Proof. If $y, y^* \varepsilon \Sigma$ satisfying those two relations, then from the second one we get $(y, y) = (y^*, y^*) = \lambda^2 \cdot (x, x)$. By computation

$$G(x, y, y^*) = -(x, x) \cdot [\lambda^2 \cdot (x, x) - (y, y^*)]^2 \leqq 0,$$

by Q_2 and Theorem 3.2. Use of S_2 gives $(y, y^*) = \lambda^2 \cdot (x, x)$, and hence $(y, y) = (y, y^*) = (y^*, y^*)$, from which $y = y^*$ follows by Q_3.

Remark. Q_3 defines an equivalence relation in $\{\Sigma : Q_1, Q_2, S_2, E_1\}$. For if $x = y$ and $y = z$ by Q_3, then $x = z$ is established by applying S_2 to $G(x. y, z)$ and using Theorems 3.1, 3.2 in case $(x, x) = 0$. The symmetric and reflexive properties are obvious.

DEFINITION 3.1. *If $x \varepsilon \Sigma$, $(x \neq \theta)$, and λ is a real number, an element y of Σ is called a scalar multiple of x by λ provided $(x, y) = \lambda \cdot (x, x)$ and $G(x, y) = 0$. We write $y = \lambda \cdot x$. If $x = \theta$, put $\lambda \cdot x = \theta$, for every λ.*

THEOREM 3.4. *For each x and each real λ, Σ contains at most one element $\lambda \cdot x$.*

Remark. For each element x *of* Σ, $0 \cdot x = \theta$.

THEOREM 3.5. *If x, $\lambda \cdot x$, $y \varepsilon \Sigma$, $(\lambda \cdot x, y) = \lambda \cdot (x, y)$.*

Proof. The theorem being obvious if $x = \theta$, suppose the contrary. Now

$$G(x, \lambda \cdot x, y) = -(x, x) \cdot [(\lambda \cdot x, y) - \lambda \cdot (x, y)]^2 \leqq 0,$$

and using S_2 gives $(\lambda \cdot x, y) = \lambda \cdot (x, y)$.

THEOREM 3.6. *If x, $1 \cdot x \varepsilon \Sigma$, then $1 \cdot x = x$.*

Proof. Use of Theorem 3.5 gives $(1 \cdot x, 1 \cdot x) = (1 \cdot x, x) = (x, x)$ and $x = 1 \cdot x$ by Q_3.

THEOREM 3.7. *If x, $\lambda_2 \cdot x$, $\lambda_1 \cdot (\lambda_2 \cdot x)$, $(\lambda_1 \lambda_2) \cdot x \varepsilon \Sigma (\lambda_1, \lambda_2$ real), then $\lambda_1 \cdot (\lambda_2 \cdot x) = (\lambda_1 \lambda_2) \cdot x$.*

Proof. The proof follows immediately from Q_2, Theorem 3.5, and Q_3.

THEOREM 3.8. *If $x, y \varepsilon \Sigma$ $(x \neq \theta)$, then y is a scalar multiple of x if and only if $G(x, y) = 0$.*

Proof. If $y = \lambda \cdot x$, then $G(x, y) = 0$ by Definition 3.1. If $G(x, y) = 0$, put $(x, y)/(x, x) = \lambda$. Then $(x, y) = \lambda \cdot (x, x)$ which, together with $G(x, y) = 0$, gives $y = \lambda \cdot x$.

THEOREM 3.9. *If* $x, y, z \in \Sigma$, $(y \neq \theta)$, *and* $G(x,y) = G(y,z) = 0$, *then* $G(x,z) = 0$.

Proof. The theorem is trivial if $x = \theta$. In the contrary case, $y = \lambda_1 \cdot x$ and $z = \lambda_2 \cdot y$ by Theorem 3.8. Then Theorem 3.7 gives $z = \lambda_2(\lambda_1 \cdot x) = (\lambda_1\lambda_2) \cdot x$, and $G(x,z) = 0$.

4. Introduction of a distance in $\{\Sigma: Q_1, Q_2, Q_3, S_1, S_2, E_1\}$. Attach to each two elements x, y of Σ the number

$$(*) \qquad xy = [(x,x) + (y,y) - 2(x,y)]^{\frac{1}{2}}$$

as distance. Clearly $xy = yx$ and $xx = 0$. Now

$$(\dagger) \quad (x,x) + (y,y) - 2(x,y) = [(x,x)^{\frac{1}{2}} - (y,y)^{\frac{1}{2}}]^2 + 2[(x,x)^{\frac{1}{2}}(y,y)^{\frac{1}{2}} - (x,y)] \geqq 0,$$

by S_1, and so xy is real and non-negative. Moreover, if $xy = 0$ then each summand in the right-hand member of $(\dagger)$ vanishes. This gives $(x,x) = (x,y) = (y,y)$ and consequently $x = y$. Thus the distance xy defined by $(*)$ makes Σ a *semimetric* space.

We note that $\theta x = (x,x)^{\frac{1}{2}}$ and $(x,y) = \frac{1}{2}(\theta x^2 + \theta y^2 - xy^2)$.

DEFINITION 4.1. *A distance space is metric about one of its elements* t *provided* t, u, v *are congruent with three points of the euclidean plane for every pair of elements* u, v *of the space.*

THEOREM 4.1. *The space* Σ *is metric about* θ.

Proof. Since Σ is semimetric, it suffices to show that the determinant $D(\theta, x, y)$ is negative or zero for every $x, y \in \Sigma$. (See [3, p. 56] for the definition of the Cayley-Menger determinant $D(p_1, p_2, \cdots, p_k)$ of a k-tuple of elements $p_1, p_2, \cdots, p_k$ of a semimetric space).

An easy computation shows that $D(\theta, x, y) = -4G(x,y)$, and hence $D(\theta, x, y) \leqq 0$ by S_1.

Remark. The space Σ is not necessarily metric. For if Σ consists of the quadruple θ, x_1, x_2, x_3 with $(\theta, \theta) = 0$, $(\theta, x_1) = (x_i, \theta) = 0$, $(x_i, x_i) = 16$, $(i = 1, 2, 3)$, and $(x_1, x_2) = (x_2, x_1) = (x_1, x_3) = (x_3, x_1) = (31)/2$, $(x_2, x_3) = (x_3, x_2) = (23)/2$, then Postulates $Q_1, Q_2, Q_3, S_1, S_2, E_1$ are satisfied. But $x_1x_2 = x_1x_3 = 1$, $x_2x_3 = 3$, and consequently the space is not metric.

DEFINITION 4.2. *Three elements* x, y, z *of* Σ *are algebraically linear provided* $B(x,y,z) = G(x,y,z) = 0$.

Remark. Since $B(x, y, z) = \frac{1}{4} \cdot D(x, y, z)$, elements x, y, z are *metrically* linear (that is, the sum of two of the distances xy, yz, xz equals the third) if and only if $B(x, y, z) = 0$. This, however, does not imply that $G(x, y, z) = 0$, and so algebraic linearity implies metric linearity, but not conversely.

5. Algebraic middle elements in$\{\mathfrak{S}: Q_1, Q_2, Q_3, S_1, S_2, E_1\}$.

DEFINITION 5.1. *If $x, z \in \mathfrak{S}$ $(x \neq z)$, an element y of $\mathfrak{S}$ is an algebraic middle element of x and z provided*

(a). *$G(x, z) \neq 0$, x, y, z are algebraically linear and $G(x, y) = G(y, z)$, or*

(b). *$G(x, z) = 0$ and $y = [(1 + \lambda)/2] \cdot x$,*

where $z = \lambda \cdot x$ (with the roles of x and z reversed in the two immediately preceding equalities in case $x = \theta$).

THEOREM 5.1. *If y is an algebraic middle element of x, z then y is a metric middle element of x and z.*

Proof. From $D(x, y, z) = 4B(x, y, z) = 0$, it follows that x, y, z are metrically linear. It remains to show that $xy = yz$.

Case 1. $G(x, z) \neq 0$. Since $G(x, y, z) = 0$, the easily established relation $D(\theta, x, y, z) = 8 \cdot G(x, y, z)$ gives $D(\theta, x, y, z) = 0$, with θ, x, y, z a metric quadruple (due to Theorem 4.1 and the metric linearity of x, y, z). Hence points θ', x', y', z' of the euclidean plane exist such that the congruence

$$(*) \qquad \theta, x, y, z \approx \theta', x', y', z',$$

holds and x', y', z' lie on a straight line [3, p. 56]. Now

$$G(x, y) = -\tfrac{1}{4}D(\theta, x, y) = -\tfrac{1}{4}D(\theta', x', y') = 4\Delta^2(\theta', x', y')$$

and, similarly, $G(y, z) = 4\Delta^2(\theta', y', z')$, where $\Delta(\ \ )$ denotes the area of the triangle whose vertices appear within the parentheses.

Then $\Delta(\theta', x', y') = \Delta(\theta', y', z')$, which follows from $G(x, y) = G(y, z)$, gives (using the linearity of x', y', z') $x'y' = y'z'$ and so $xy = yz$ from congruence (*).

Case 2. $G(x, z) = 0$. If $x \neq \theta$, then $y = [(1 + \lambda)/2] \cdot x$ and

$$xy = \{(x, x) + [(1 + \lambda)/2]^2 \cdot (x, x) - (\lambda + 1) \cdot (x, x)\}^{\frac{1}{2}};$$

that is, $xy = \frac{1}{2} \mid 1 - \lambda \mid \cdot (x, x)^{\frac{1}{2}}$. A similar computation shows that yz has

the same value. If, finally, $x = \theta$, then $y = (1/2) \cdot z$ and we get $xy = \theta y = (1/2) \cdot (z, z)^{\frac{1}{2}} = yz$.

Though the determinant D formed for a finite subset of Σ is evidently a congruence invariant, the Gram determinants G do not have this property. If, however, $x_1, x_2, \cdots, x_n$ and $y_1, y_2, \cdots, y_n$ are two n-tuples of Σ such that $x_1, x_2, \cdots, x_n \approx y_1, y_2, \cdots, y_n$ and $\theta x_i = \theta y_i$, $(i = 1, 2, \cdots, n)$, then $G(x_1, x_2, \cdots, x_n) = G(y_1, y_2, \cdots, y_n)$, since each is equal to

$$(-1)^{n+1} \cdot 2^{-n} D(\theta, x_1, x_2, \cdots, x_n).$$

LEMMA 5.1. *If y is an algebraic middle element of x and z then*

$$\theta y = (1/2)[2(\theta x^2 + \theta z^2) - xz^2]^{\frac{1}{2}}.$$

Proof. Developing $D(\theta, x, y, z) = 8G(x, y, z) = 0$, and using $xy = yz = (1/2)xz$ (Theorem 5.1) the expression for θy given in the lemma is obtained.

6. The system $\{\boldsymbol{\Sigma}: Q_1, Q_2, Q_3, S_1, S_2, S_3, E_1\}$. Adjoining the third Schwarz postulate to the preceding system, we prove the following theorem.

THEOREM 6.1. *There is at most one algebraic middle element of two elements of Σ.*

Proof. If y, y^* are algebraic middle elements of x and z, Theorem 5.1 and Lemma 5.1 give

$$xy = yz = (1/2) \cdot xz = xy^* = y^*z, \quad \theta y = \theta y^*.$$

Case 1. $G(x, z) \neq 0$. From

$$D(\theta, x, y, z, y^*) = -16G(x, y, z, y^*) \leqq 0$$

we readily obtain $-D(\theta, x, z)(yy^*)^2 \leqq 0$. But $G(x, z) \neq 0$ gives $D(\theta, x, z) = -4G(x, z) < 0$, and it follows that $yy^* = 0$. Since Σ is semimetric, this implies $y = y^*$, and the theorem is proved in this case.

Case 2. $G(x, z) = 0$. If $x \neq \theta$ then $y = [(1 + \lambda)/2] \cdot x$, where $z = \lambda \cdot x$, and the uniqueness of y follows from Theorem 3.4. A similar argument is employed in case $x = \theta$ (and consequently $z \neq \theta$), and the proof is complete.

DEFINITION 6.1. *Let x, z be elements of Σ with algebraic middle element y. An element $2 \cdot y$ is called a sum of x and z (written $2 \cdot y = x + z$). If $x = z$, put $x + z = 2 \cdot x$.*

If $2 \cdot y = x + z$ then $(1/2)(2 \cdot y) = (1/2)(x + z)$ or (Theorem 3.7) $y = \frac{1}{2}(x + z)$.

THEOREM 6.2. *Two elements x, z of Σ have at most one sum $x + z$.*

LEMMA 6.1. *If $x, y, x + y$ are elements of Σ, then*

$$(x + y, x) = (x, x) + (x, y), \quad (x + y, x + y) = (x, x) + (y, y) + 2(x, y).$$

Proof. Writing $2 \cdot w = x + y$, we have

$$(x + y, x) = (2 \cdot w, x) = 2 \cdot (w, x) = \theta w^2 + \theta x^2 - x w^2.$$

Since $x w^2 = (1/4) x y^2$ and (Lemma 5.1) $\theta w^2 = \frac{1}{2}(\theta x^2 + \theta y^2) - (1/4) \cdot x y^2$, substitution yields

$$(x + y, x) = \theta x^2 + (1/2)(\theta x^2 + \theta y^2 - x y^2) = (x, x) + (x, y).$$

Also,

$$\begin{aligned}(x + y, x + y) &= (2 \cdot w, 2 \cdot w) = 4 \cdot (w, w) = 4 \cdot \theta w^2 = 2(\theta x^2 + \theta y^2) - x y^2 \\ &= 2(\theta x^2 + \theta y^2) - (\theta x^2 + \theta y^2 - 2(x, y)) = \theta x^2 + \theta y^2 + 2 \cdot (x, y) \\ &= (x, x) + (y, y) + 2 \cdot (x, y).\end{aligned}$$

THEOREM 6.2. *If $x, y, x + y, z \in \Sigma$, then $(x + y, z) = (x, z) + (y, z)$.*

Proof. Suppose, first, that $G(x, y) \neq 0$. Since

$$G(x, y, x + y) = 4 \cdot G(x, y, \tfrac{1}{2}(x + y)) = 0,$$

S_3 yields $G(x, y, x + y, z) \geqq 0$.

Use of Lemma 6.1 gives

$$G(x, y, x + y, z) = -[(x + y, z) - (x, z) - (y, z)]^2 \cdot G(x, y) \leqq 0.$$

Consequently $G(x, y, x + y, z) = 0$ and $(x + y, z) = (x, z) + (y, z)$.

If $G(x, z) = 0$, then (assuming $x \neq \theta$), $\frac{1}{2}(x + y) = [(1 + \lambda)/2] \cdot x$ with $y = \lambda \cdot x$, and so $x + y = (1 + \lambda) \cdot x$ (Theorem 3.7). Whence

$$\begin{aligned}(x + y, z) &= ((1 + \lambda) \cdot x, z) = (1 + \lambda) \cdot (x, z) \\ &= (x, z) + (\lambda \cdot x, z) = (x, z) + (y, z).\end{aligned}$$

The desired result is obtained in a similar manner in case $x = \theta$ (and consequently $y \neq \theta$).

7. The system $\{\mathfrak{S}: Q_1, Q_2, Q_3, S_1, S_2, S_3, E_1, E_2, E_3\}$. We adjoin now the two existence postulates E_2 and E_3 to our system.

THEOREM 7.1. *For each element x of $\mathfrak{S}$ and each real number λ, $\mathfrak{S}$ contains exactly one element $\lambda \cdot x$. If $x, y \in \mathfrak{S}$ there is exactly one element $x + y$ of $\mathfrak{S}$.*

Proof. The proof of the first statement of the theorem is an immediate consequence of Postulate E_2, Definition 3.1, and Theorem 3.4, while the proof of the second statement follows from the first one and Theorem 6.1, when it is shown that at least one algebraic middle element exists for every pair of distinct elements of $\mathfrak{S}$. (If $x = z$, $x + z = 2 \cdot x$ exists by the first part of the theorem).

If $x, z \in \mathfrak{S}$ $(x \neq z)$, Postulate E_3 insures the existence of at least one middle element of x, z whenever $G(x, z) \neq 0$. If x, $z \in \mathfrak{S}$, $(x \neq z)$ with $G(x, z) = 0$, then $z = \lambda \cdot x$ (assuming $x \neq \theta$) and the element $[(1 + \lambda)/2] \cdot x$ which exists by the first part of the theorem, is the desired algebraic middle element of x and z. If, finally, $x = \theta$ then the element $(1/2) \cdot z$ is an algebraic middle element of x and z.

THEOREM 7.2. *Addition and scalar multiplication in $\mathfrak{S}$ have the following properties:*

(i) $x + y = y + x$,

(ii) $x + (y + z) = (x + y) + z$,

(iii) $x + y = x + z$ *implies* $y = z$,

(iv) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$,

(v) $(\lambda_1 + \lambda_2) \cdot x = \lambda_1 \cdot x + \lambda_2 \cdot x$.

Proof. Properties (i), (ii), (iv), (v) follow at once from Postulates Q_1, Q_3, and Theorems 3.5, 6.2. To establish (iii), we have $x + y = x + z$ implies $(x + y, y) = (x + z, y)$, and consequently $(y, y) = (y, z)$. Also, $(x + y, z) = (x + z, z)$ and so $(y, z) = (z, z)$. Hence $y = z$ by Q_3.

THEOREM 7.3. *Distances θx, $x \in \mathfrak{S}$, have the following properties:*

(a) $\theta x = (x, x)^{\frac{1}{2}} \geqq 0$,

(b) $\theta x = 0$ *if and only if* $x = \theta$,

(c) $$\text{dist. } [\theta, \lambda \cdot x] = |\lambda| \cdot \theta x,$$

(d) $$\text{dist. } [\theta, x + y] \leqq \theta x + \theta y.$$

Proof. The relation $\theta x = (x, x)^{\frac{1}{2}}$ is given in the second paragraph of Section 4, and (b) follows from Q_2 and Theorem 3. 2.

Since

$$\begin{aligned}\text{dist. } [\theta, \lambda \cdot x] &= [(\theta, \theta) + (\lambda \cdot x, \lambda \cdot x) - 2(\theta, \lambda \cdot x)]^{\frac{1}{2}} \\ &= (\lambda \cdot x, \lambda \cdot x)^{\frac{1}{2}} = |\lambda| \cdot (x, x)^{\frac{1}{2}} = |\lambda| \cdot \theta x,\end{aligned}$$

property (c) is proved.

To establish (d) we have from $G(x, y) \geqq 0$,

$$[(x, x)^{\frac{1}{2}} + (y, y)^{\frac{1}{2}}]^2 \geqq (x, x) + (y, y) + 2\,|\,(x, y)\,|\,,$$

or

(*) $$\theta x + \theta y \geqq (x + y, x + y)^{\frac{1}{2}} = \text{dist. } [\theta, x + y].$$

We are justified by Theorem 7. 3 in defining the *norm* $\| x \|$ of an element x of $\mathfrak{Z}$ as the distance θx. Noting that

$$\begin{aligned}(x - y, x - y)^{\frac{1}{2}} &= (x + (-1)y, x + (-1)y)^{\frac{1}{2}} \\ &= [(x, x) + (y, y) - 2(x, y)]^{\frac{1}{2}} = xy,\end{aligned}$$

or $xy = \| y - x \| = \text{dist. } [\theta, y - x]$, it follows that for each three elements x, y, z of $\mathfrak{Z}$

$$\text{dist. } [y - x, z - x] = \| (z - x) - (y - x) \| = \| z - y \| = yz.$$

Remark. The space $\mathfrak{Z}$ is metric.

THEOREM 7. 4. *The space* $\{\mathfrak{Z}: Q_1, Q_2, Q_3, S_1, S_2, S_3, E_1, E_2, E_3\}$ *is a generalized euclidean space.*

Proof. Postulates Q_1, Q_2 and Theorems 3. 2, 3. 5, 6. 2 prove that the quasi inner product (x, y) has the properties (a)-(e) of an inner product (Section 2), and Theorems 3. 6, 3. 7, 7. 2) establish that $\mathfrak{Z}$ is a linear space with respect to the addition $x + y$ and the scalar multiplication $\lambda \cdot x$.

By Theorem 7. 3 the space $\mathfrak{Z}$ is normed by taking $\| x \| = \theta x$, and since $\theta x = (x, x)^{\frac{1}{2}}$, $\mathfrak{Z}$ satisfies all the requirements of Definition 2. 1 for a generalized euclidean space.

8. Some further properties of $\mathfrak{S}$. We understand now by $\mathfrak{S}$ a set subjected to all nine of the postulates in Section 2; that is, $\mathfrak{S}$ is a generalized euclidean space (Theorem 7.4). It is well-known that if such a space is separable and complete, it is congruent with a euclidean space or with Hilbert space according as its dimension is finite or infinite, respectively. It might be of interest in this concluding section to give another proof of this important theorem on the basis of the theory we have been developing.

THEOREM 8.1. *Metric linearity and algebraic linearity are equivalent in* $\mathfrak{S}$.

Proof. By the Remark following Definition 4.2, algebraic linearity implies metric linearity. To prove the converse it suffices to show that if $x, y, z \in \mathfrak{S}$, then $B(x, y, z) = 0$ implies $G(x, y, z) = 0$. This is obvious in case the elements are not pairwise distinct, or if one of the elements is θ. Suppose neither alternative holds.

By the concluding remarks of Section 7, it is easily seen that $x, y, z \approx \theta, y - x, z - x$, and so the metric linearity of x, y, z implies that of $\theta, y - x, z - x$. Consequently $G(y - x, z - x) = -(1/4)D(\theta, y - x, z - x) = 0$, and application of Theorem 3.8 gives

$$z - x = \lambda \cdot (y - x); \quad z = (1 - \lambda) \cdot x + \lambda \cdot y,$$

from which $G(x, y, z) = 0$ follows by computation.

DEFINITION 8.1. *A space has the weak euclidean four-point property provided each quadruple of its points containing a linear triple is congruent with four points of the euclidean plane.*

THEOREM 8.2. *The space* $\mathfrak{S}$ *has the weak euclidean four-point property.*

Proof. Let x, y, z, w be any four points of $\mathfrak{S}$ (pair-wise distinct) with x, y, z metrically linear. Since $x, y, z, w \approx x - w, y - w, z - w, \theta$, it suffices to prove the latter quadruple imbeddable congruently in the plane. The space $\mathfrak{S}$ being metric, this follows upon showing $D(\theta, x - w, y - w, z - w) = 0$, [3].

Now the metric linearity of x, y, z implies that of $x - w$, $y - w$, $z - w$, and hence (Theorem 8.1) this triple is algebraically linear; that is $G(x - w, y - w, z - w) = 0$. But

$$D(\theta, x - w, y - w, z - w) = 8 \cdot G(x - w, y - w, z - w) = 0,$$

and the proof is complete.

Since for each two of its distinct elements x, z the space Σ contains an algebraic (and hence metric) middle element, Σ is *metrically convex.* If, moreover, $x, y \in \Sigma$ the element $z = 2y - x$ is such that y is the middle element of x and z, and so Σ is *externally convex.*

DEFINITION 8.2. *An element x of Σ is a limit of an infinite sequence $\{x_n\}$ of elements of Σ provided* $\lim(x, x_n) = \lim(x_n, x_n) = (x, x)$.

Remark. Writing $(x, x) + (x_n, x_n) - 2(x, x_n)$ as in (†) of Section 4, it is seen that $\lim x_n = x$ if and only if $\lim xx_n = 0$.

Cauchy sequences and separability are now easily defined in terms of inner product, and we have the following theorems.

THEOREM 8.3. *The space Σ is congruent with Hilbert space provided it is separable, complete, and for each positive integer k a subset $x_1, x_2, \cdots, x_k$ of Σ exists with $G(x_1, x_2, \cdots, x_k) \neq 0$.*

THEOREM 8.4. *The space Σ is congruent with the n-dimensional euclidean space E_n provided it is complete, and n is the smallest positive integer such that each $n + 1$ elements of Σ have a vanishing determinant G.*

Since Σ is metrically convex, externally convex, and has the weak euclidean four-point property, these theorems are immediate consequences of metric characterization theorems of Hilbert and euclidean spaces (with respect to the class of all semimetric spaces), respectively [3, p. 70], observing (as noted in Section 5) that for each positive integer k,

$$D(\theta, x_1, x_2, \cdots, x_k) = (-1)^{k+1} \cdot 2^k G(x_1, x_2, \cdots, x_k).$$

It is observed that a congruence of Σ with E_n or with Hilbert space preserves inner products if and only if the element θ' corresponding to θ in the congruence is the origin (or null vector) of the space. Since this can always be brought about by a translation, Theorems 8.3 and 8.4 can be restated in terms of a *vectorial application* of the vector space Σ onto the euclidean or Hilbert vector spaces.

In conclusion, we wish to point again to the fact that we have avoided the obvious and ad hoc procedure that, for example, defines an element c as a sum of elements a, b provided $(c, x) = (a, x) + (b, x)$ *for every element x of Σ.* Such a device places a condition on the inner products of a, b, c with *every* element of the space in order to ascertain if one of the three elements is a sum of the other two! This question is answered, according to the procedure of the paper, merely by examining the inner products

attached to a, b, and c, the three elements concerned. The linearity of inner product clearly partakes more of the nature of a theorem than a postulate, and much of the paper is concerned with establishing that property from our (a priori) milder assumptions.

THE UNIVERSITY OF MISSOURI.

REFERENCES.

[1] N. Aronszajn, "Caractérisation métrique de l'espace de Hilbert, des espaces vectoriels et de certains groupes métriques," *Comptes Rendus, Academie des Sciences*, vol. 201 (1935), pp. 811-813; pp. 873-875.

[2] S. Banach, *Théorie des operations linéaires*, Warsaw, 1932.

[3] L. M. Blumenthal, *Distance Geometries*, University of Missouri Studies, vol. 13, 1938.

[4] M. Fréchet, "Sur la définition axiomatique d'une classe d'espaces vectoriels distanciés applicables vectoriellement sur l'espace de Hilbert," *Annals of Mathematics*, vol. 36 (1935), pp. 705-718.

[5] P. Jordan and J. von Neumann, "On inner products in linear metric spaces," *ibid.*, vol. 36 (1935), pp. 719-723.

[6] K. Menger, "Neuer Aufbau der Vektoralgebra," *Ergebnisse eines mathematischen Kolloquiums* (Wien), Heft 5 (1933), pp. 27-29; *L'Enseignement Mathématique*, nos. 5-6 (1936), pp. 359-362.

ON THE SPECTRUM OF HILBERT'S MATRIX.* [1]

By WILHELM MAGNUS.

1. Introduction and summary. It has been shown by Hilbert that the infinite matrix

$$(1.1) \qquad (\alpha_{nm}), \text{ where } \alpha_{nm} = (n+m+1)^{-1}; \qquad n, m = 0, 1, 2, \cdots,$$

is bounded, and Hilbert's inequality,

$$(1.2) \qquad 0 \leqq \sum_{n=0}^{\infty} \sum_{m=0}^{\infty} x_n x_m/(n+m+1) \leqq \pi \sum_{n=0}^{\infty} x^2_n,$$

for real x_n, has been proved in many different ways; for references cf. Hardy, Littlewood and Pólya [2]. The finite segments of (1.1) have been investigated by H. Frazer [1] and O. Taussky [6]. The matrix (1.1) is associated with Legendre's polynomials (cf. Szegö [4]), and the quadratic form in (1.2) has been used by Szegö [5] for a normalization of Hankel forms.

Hilbert's original proof of (1.2) (cf. H. Weyl [7], I. Schur [3]) suggests that the spectrum of (1.1) is purely continuous. We shall obtain the following result: [2]

THEOREM 1. *The spectrum of Hilbert's matrix* (1.1) *is purely continuous. Every real value of* λ *for which* $0 \leqq \lambda \leqq \pi$ *belongs to the spectrum.*

2. An integral equation. We shall prove

THEOREM 2. *If the real numbers* g_n, h_n, $(n = 0, 1, 2, \cdots)$ *satisfy*

$$(2.1) \qquad \sum_0^{\infty} g^2_n = M < \infty, \qquad \sum_0^{\infty} h^2_n = N < \infty$$

---

* Received January 25, 1950.

[1] Work sponsored by the Office of Naval Research.

[2] I am indebted to Professor Wintner for calling my attention to the fact that the second statement of Theorem 1 can be derived from Carleman's work (*Sur les équations intégrales singulières à noyau réel et symmétrique*, Uppsala, 1923, pp. 169) on the integral analogue of Hilbert's matrix. Carleman shows that the kernel $(s+t)^{-1}$, $0 \leqq s, t < \infty$, has the purely continuous spectrum $[0, \pi]$; from this, from Hilbert's original proof of (1.2), and from a result due to H. Weyl (*Rendiconti del Circolo Matematico di Palermo*, vol. 27 (1909), pp. 373-392), it follows that every point of $[0, \pi]$ belongs to the spectrum of (1.1).

(which implies that the power series

$$g(z) = \sum_0^\infty g_n z^n, \qquad h(z) = \sum_0^\infty h_n z^n \tag{2.2}$$

converge for $|z| < 1$*), and if* λ *denotes a real number such that* $0 \leqq \lambda \leqq \pi$*, then neither of the integral equations*

$$\int_0^1 g(t)/(1-tz)\,dt = \lambda g(z), \tag{2.3}$$

$$\int_0^1 h(t)/(1-tz)\,dt = \lambda h(z) + 1 \tag{2.4}$$

has a solution, except for the solution $g(z) \equiv 0$ *of* (2.3).

This is equivalent to Theorem 1. We can show this by expanding $(1-tz)^{-1}$ in a series of ascending powers of zt, introducing this expansion into (2.3) and (2.4), and observing that, because of (2.1), term by term integration is permitted.

To prove Theorem 2, we deduce first from (2.1) and from Schwarz's inequality that, for $|z| < 1$,

$$|g(z)| \leqq M^{\frac{1}{2}} |1-z^2|^{-\frac{1}{2}}, \qquad |h(z)| \leqq N^{\frac{1}{2}} |1-z^2|^{-\frac{1}{2}}. \tag{2.5}$$

This shows that we can iterate (2.3) and (2.4) which gives

$$\int_0^1 \{g(t)/(t-z)\}\log\{(1-z)/(1-t)\}\,dt = \lambda^2 g(z), \tag{2.6}$$

$$\int_0^1 \{h(t)/(t-z)\}\log\{(1-z)/(1-t)\}\,dt = \lambda^2 h(z) + \lambda - z^{-1}\log(1-z). \tag{2.7}$$

We now introduce

$$1-t = e^{-\tau}, \qquad 1-z = e^{-\zeta}, \qquad g(t) = \gamma(\tau), \qquad h(t) = \eta(\tau). \tag{2.8}$$

Multiplying (2.6) and (2.7) by $e^{-p\zeta}$ and integrating with respect to ζ from 0 to ∞, we find

$$\begin{aligned} \lambda^2 \int_0^\infty \gamma(\zeta) e^{-p\zeta} d\zeta \\ = \int_0^\infty \gamma(\tau) e^{-p\tau} d\tau \{ \int_0^\infty e^{-p\sigma} \sigma/(1-e^{-\sigma})\,d\sigma \\ + \int_0^\infty e^{-\sigma(1-p)} \sigma\, d\sigma/(1-e^{-\sigma}) - \int_\tau^\infty e^{-\sigma(1-p)} \sigma/(1-e^{-\sigma})\,d\sigma \}. \end{aligned} \tag{2.9}$$

The exchanging of the integrations is permitted if $\operatorname{Re} p > \frac{1}{2}$, where Re denotes the real part.

The integrals from 0 to ∞ with respect to σ in (2. 9) can be evaluated explicitly, and the integral from τ to ∞ can be evaluated by an expansion of $\{1 - \exp(-\sigma)\}^{-1}$ into a sum of terms $\exp(-n\sigma)$, provided that $\operatorname{Re} p < 1$. The result is

$$(2.10)\quad \{(\pi/\sin \pi p)^2 - \lambda^2\} \int_0^\infty e^{-p\zeta}\gamma(\zeta)\,d\zeta = \sum_{k=0}^{\infty} \{b_k(1-p+k)^{-1} + a_k(1-p+k)^{-2}\},$$

where $\frac{1}{2} < \operatorname{Re} p < 1$ and

$$(2.11)\quad a_k = \int_0^\infty \gamma(\sigma)e^{-(k+1)\sigma}d\sigma = \int_0^1 g(t)\,(1-t)^k dt,$$

$$(2.12)\quad b_k = \int_0^\infty \gamma(\sigma)e^{-(k+1)\sigma}d\sigma = -\int_0^1 g(t)\,(1-t)^k \log(1-t)\,dt.$$

But since the integral on the left-hand side of (2. 9) is an analytic function of p if $\operatorname{Re} p > \frac{1}{2}$, the latter condition is the only restriction for the validity of (2. 10), the right hand side being regular in p except for poles at $p = k + i$ $(k = 0, 1, 2, \cdots)$. Since the left-hand side in (2. 10) vanishes at

$$(2.13)\quad p = \tfrac{1}{2} + n \pm i\theta, \qquad (\theta = \pi^{-1}\{\log(\pi + (\pi^2 - \lambda^2)^{\frac{1}{2}}) - \log \lambda\}),$$

where $n = 1, 2, \cdots$, the same must hold for the right-hand side of (2. 10).

We can deduce from (2. 1) that an application of the inversion formula of the Laplace transform

$$(2.14)\quad \int_0^\infty e^{-p\zeta}\gamma(\zeta)\,d\zeta$$

must give $\gamma(\zeta)$. If we compute (2. 14) from (2. 10), this supplies a representation of $\gamma(\zeta)$ by an integral taken over a parallel to the imaginary axis of the p-plane, which then can be evaluted by an application of the formula of residues. To prove this, we deduce from (2. 5) that

$$(2.15)\quad |a_k| \leqq M^{\frac{1}{2}}(k+\tfrac{1}{2})^{-1}, \qquad |b_k| \leqq M^{\frac{1}{2}}(k+\tfrac{1}{2})^{-2},$$

which shows that the series on the right-hand side of (2. 10) tends to zero if $\operatorname{Re} p < 1$ and $|p| \to \infty$. Going back from ζ to z, the result is, for $0 < \lambda < \pi$,

$$(2.16)\quad g(z) = \tfrac{1}{2}\lambda^{-2}(\pi^2 - \lambda^2)^{-\frac{1}{2}} \sum_{m=0}^{\infty} (1-z)^{m-\frac{1}{2}}\{q_m(1-z)^{-i\theta} + \bar{q}_m(1-z)^{i\theta}\},$$

where θ is defined by (2.13) and q_n by

$$(2.17)\quad q_m = -i\sum_{k=0}^{\infty}\{a_k(\tfrac{1}{2}+m+k-i\theta)^{-2}+b_k(\tfrac{1}{2}+m+k-i\theta)^{-1}\}.$$

In a similar way, we find from (2.7) that

$$(2.18)\quad h(z) = \tfrac{1}{2}\lambda^{-2}(\pi^2-\lambda^2)^{-\frac{1}{2}}\sum_{m=0}^{\infty}(1-z)^{m-\frac{1}{2}}\{r_m(1-z)^{-i\theta}+\bar{r}_m(1-z)^{i\theta}\},$$

where

$$(2.19)\quad r_m+\lambda(\tfrac{1}{2}-m+i\theta)^{-1}+i\sum_{k=0}^{\infty}(\tfrac{1}{2}+k-m+i\theta)^{-2}$$
$$= -i\sum_{k=0}^{\infty}\{\alpha_k(\tfrac{1}{2}+m+k-i\theta)^{-2}+\beta_k(\tfrac{1}{2}+m+k-\theta)^{-1}\},$$

$$(2.20)\quad \alpha_k = \int_0^1 h(t)(1-t)^k dt, \qquad \beta_k = -\int_0^1 h(t)(1-t)^k \log(1-t)dt.$$

Finally, it follows from (2.3), by Stieltjes' inversion formula, that, if $w>1$,

$$(2.21)\quad \lim_{\epsilon\to 0}\{g(w+i\epsilon)-g(w-i\epsilon)\} = 2\pi i(w/\lambda)g(w^{-1}),$$

where $g(w\pm i\epsilon)$ denotes the analytic continuation of $g(z)$ into the neighborhood of the upper and lower side of the real axis. (2.21) is also satisfied if we substitute $h(z)$ for $g(z)$. Combined with (2.16), this gives

$$(2.22)\quad w\sum_{m=0}^{\infty}(-1)^m(w-1)^{m-\frac{1}{2}}\{q_m(w-1)^{-i\theta}+\bar{q}_m(w-1)^{i\theta}\}$$
$$= \sum_{m=0}^{\infty}(1-w^{-1})^{m-\frac{1}{2}}\{q_m(1-w^{-1})^{-i\theta}+\bar{q}_m(1-w^{-1})^{i\theta}\}.$$

From this we see, by letting $w\to 1$, that if $q_0=0$, then $q_1=0$. But q_0 must be zero, because

$$(2.33)\quad \int_0^1\{g(z)\}^2dz = \sum_{n,m=0}^{\infty} g_n g_m(n+m+1)^{-1} \leqq \pi\sum_0^{\infty} g^2{}_n$$

is finite, according to (1.2) and (2.1). On the other hand, it follows from (2.16) that the integral on the left in (2.23) is infinite unless $q_0=0$. Hence, if $g(z)$ satisfies (2.1) and (2.3), and if $0<\lambda<\pi$, then

$$(2.24)\quad g(1)=g'(1)=0, \qquad |g'(z)|\leqq M',$$

where M' is a constant and $0\leqq z\leqq 1$. Similarly, we find

$$(2.25)\quad h(1)=h'(1)=0, \qquad |h'(z)|\leqq N'.$$

Differentiating (2.3) and (2.4) with respect to z, we have from (2.24), (2.25)

$$\lambda g'(z) = \int_0^1 (1-tz)^{-2}\{tg(t)\}dt$$

$$= -z^{-1}\int_0^1 (1-zt)^{-1}\{tg'(t)+g(t)\}dt, \tag{2.26}$$

and (2.26) is also valid if we substitute $h(z)$ for $g(z)$.

If we combine (2.26) and (2.3), we obtain

$$\lambda\{zg'(z)+g(z)\} = \int_0^1 tg'(t)(1-tz)^{-1}dt.$$

Since $|g'(t)| \leqq M'$, this gives $\lambda(n+1)|g_n| \leqq M'\int_0^1 t^{n+1}dt$, which shows that

$$\sum_0^\infty \{(n+1)g_n\}^2 < \infty. \tag{2.27}$$

We can replace the second equation in (2.26) by an infinite system of linear equations, viz.,

$$\sum_{m=0}^\infty (m+1)g_m(n+m+1)^{-1} = -\lambda n g_n, \qquad (n=0,1,2,\cdots).$$

Putting $(m+1)g_m = x_m$, we find

$$\sum_{n=0}^\infty \sum_{m=0}^\infty x_n x_m/(n+m+1) = -\lambda \sum_{n=0}^\infty n/(n+1)x^2{}_n. \tag{2.28}$$

Because of (2.27), we can apply (1.2), obtaining a contradiction unless $x_1 = x_2 = \cdots = 0$. Hence we see from (2.28) that also $x_0 = 0$, and therefore $g(z) \equiv 0$. The proof for $h(z) \equiv 0$ (i.e., for the non-existence of $h(z)$) is precisely the same. This proves Theorem 2 for $0 < \lambda < \pi$.

To complete the proof of Theorem 2, it will be sufficient to show that $g(z) \equiv 0$ if $\lambda = 0$ or $\lambda = \pi$, since the spectrum is a closed set. Therefore we have to exclude only the possibility that $\lambda = 0$ or $\lambda = \pi$ belongs to the point spectrum of (1.1). The case $\lambda = 0$ can easily be excluded by introducing $G(z) = \sum_0^\infty g_n z^{n+1}(n+1)^{-1}$. This is a continuous function of z for $0 \leqq z \leqq 1$, and from (2.3) we have in the case $\lambda = 0$

$$\int_0^1 g(t)t^n dt = G(1) - n\int_0^1 G(t)t^{n-1}dt = 0$$

for $n = 0, 1, 2, \cdots$. For $n = 0$, this gives $G(1) = 0$, and for $n = 1, 2, \cdots$ we find that all the moments of $G(t)$ vanish, i. e., $G(t) \equiv 0$. The case $\lambda = \pi$ can be dealt with in the same way as the general case if we modify (2. 16) according to the fact that the evaluation of $\gamma(\zeta)$ from (2. 10) requires the computation of residues at poles of the second order of the integrand. The analogue of (2. 16) then involes a term with $\log(1 - z)$.

CALIFORNIA INSTITUTE OF TECHNOLOGY.

REFERENCES.

[1] H. Frazer, "Note on Hilbert's inequality," *Journal of the London Mathematical Society*, vol. 21 (1946), pp. 7-9.

[2] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities*, Cambridge (1934), pp. 226-227.

[3] I. Schur, "Bemerkungen zur Theorie der beschränkten Bilinearformen mit unendlich vielen Veränderlichen," *Journal für die reine und angewandte Mathematik*, vol. 140 (1911), pp. 1-28.

[4] G. Szegö, *Orthogonal polynomials*. New York (1939).

[5] ———, "A Hankel-Féle Formákról," *Mathematikai es Természettudományi értesitö*, vol. 36 (1918), pp. 497-538. Review in *Jahrbuch über die Fortschritte der Mathematik*, vol. 46 (1916-1918), p. 649.

[6] O. Taussky, "A remark concerning the characteristic roots of the finite segments of the Hilbert matrix," *Quarterly Journal of Mathematics*, Oxford ser., vol. 20 (1949), pp. 80-83.

[7] H. Weyl, *Singuläre Integralgleichungen mit besonderer Berücksichtigung des Fourierschen Integraltheorems*. Dissertation, Göttingen, 1908, pp. 83-86.

THE EXPANSION THEOREM FOR SIGMA-MONOGENIC FUNCTIONS.*

By LIPMAN BERS.

1. Introduction. A Σ-monogenic function $f(z)$ is a complex-valued function $f = u + iv$ whose real and imaginary part satisfy the partial differential equations

$$\sigma_1(x)u_x = \tau_1(y)v_y, \qquad \sigma_2(x)u_y = -\tau_2(y)v_x, \tag{1.1}$$

[1, 2, 3].[1] In this note we assume that the functions σ_1, τ_1, σ_2, τ_2 are of class C'' and positive, and we say that f is Σ-monogenic only if u and v are of class C''. Without essential loss of generality we may assume that the functions σ, τ are defined for all real values of their arguments. If u and v are of class C'' and satisfy instead of (1.1) the system

$$u_x/\sigma_2(x) = \tau_1(y)v_y, \qquad u_y/\sigma_1(x) = -\tau_2(y)v_x, \tag{1.1'}$$

we say that $f = u + iv$ is Σ'-monogenic.[2]

The Σ-integral of a continuous function $f = u + iv$ is defined by the relation

$$\int f d_\Sigma z = \int u\sigma_2 dx - v\tau_2 dy + i\int (v/\sigma_1)dx + (u/\tau_1)dy. \tag{1.2}$$

It is path-independent if $f(z)$ is Σ-monogenic and is in this case a Σ'-monogenic function of the upper (variable) limit of integration. The Σ-derivative [3] of a Σ-monogenic function $f(z)$ is defined by

$$d_\Sigma f/d_\Sigma z = f^{[1]}(z) = \sigma_1 u_x + iv_x/\sigma_2 = \tau_1 v_y - iu_y/\tau_2. \tag{1.3}$$

The real and imaginary parts of $f^{[1]}$ satisfy (1.1'), so that if they are of class C'', $f^{[1]}$ is Σ'-monogenic. The higher Σ, Σ'-derivatives of the Σ-monogenic function $f(z)$, provided they exist, are defined by the recurrence formulae $f^{[0]} = f$, $f^{[2n+1]} = d_\Sigma f^{[2n]}/d_\Sigma z$, $f^{[2n]} = d_{\Sigma'} f^{[2n-1]}/d_{\Sigma'} z$.

A significant class of Σ-monogenic and Σ'-monogenic functions are the

* Received January 6, 1950; revised March 30, 1950. Presented to the Society on August 30, 1949. The research leading to this paper was supported by the ONR (Contract N6-onr-248).

[1] Numbers in brackets refer to the references at the end of the paper.

[2] The letters Σ, Σ' denote the coefficient-arrays of the systems (1.1), (1.1'). Note that $\Sigma'' = \Sigma$.

[3] Σ-integration and Σ-differentiation are inverse processes.

"formal powers" defined (for any constant values of a and z_0) by the recurrence formulae[4]

$$a \cdot Z^{(0)}(z_0; z) = a \cdot \tilde{Z}^{(0)}(z_0; z) = a,$$

$$a \cdot Z^{(n)}(z_0; z) = n \int_{z_0}^{z} a \cdot \tilde{Z}^{(n-1)}(z_0; \zeta)\, d_{\Sigma'}\zeta,$$

$$a \cdot \tilde{Z}^{(n)}(z_0; z) = n \int_{z_0}^{z} a \cdot Z^{(n-1)}(z_0; \zeta)\, d_{\Sigma}\zeta.$$

The following facts are immediate consequences of this definition. For a real λ and complex a, b

$$\lambda\{a \cdot Z^{(n)}(z_0; z)\} = (\lambda a) \cdot Z^{(n)}(z_0; z), \tag{1.4}$$

$$a \cdot Z^{(n)}(z_0; z) + b \cdot Z^{(n)}(z_0; z) = (a + b) \cdot Z^{(n)}(z_0; z). \tag{1.5}$$

A "formal polynomial"

$$P(z) = \sum_{\nu=0}^{N} a_\nu \cdot Z^{(\nu)}(z_0; z) \tag{1.6}$$

is a Σ-monogenic function possessing Σ, Σ'-derivatives of all orders. In particular,

$$P^{[\nu]}(z_0) = \nu!\, a_\nu, \qquad \nu = 0, 1, \cdots, N;\ P^{[\mu]}(z) \equiv 0, \mu > N. \tag{1.7}$$

The formal powers satisfy the inequality

$$| a \cdot Z^{(n)}(z_0; z)| \leqq | a | \{C(| z_0 | + | z |)| z - z_0 |\}^n, \tag{1.8}$$

where $C(r)$ is a positive continuous function depending upon the functions σ, τ [1, p. 76]. This implies that if $| a_\nu |^{1/\nu} = O(1)$, $\nu \to \infty$, then the "formal power series"

$$f(z) = \sum_{\nu=0}^{\infty} a_\nu \cdot Z^{(\nu)}(z_0; z) \tag{1.9}$$

converges uniformly and absolutely in some neighborhood G of z_0, and so do all series obtained from it by Σ, Σ'-differentiation. Hence $f(z)$ is Σ-monogenic in G, possesses Σ, Σ'-derivatives of all orders, and

$$f^{[\nu]}(z_0) = \nu!\, a_\nu, \qquad \nu = 0, 1, \cdots. \tag{1.10}$$

Now let $f(z)$ be a given Σ-monogenic function defined in the neighborhood of z_0. Can it be expanded in a formal power series of the form (1.9)? An affirmative answer was given (by A. Gelbart and the author) under the

[4] Note that $a \cdot Z^{(n)}$ is one symbol. In general, $aZ^{(n)} \neq a \cdot Z^{(n)}$.

hypothesis that the coefficients σ, τ are *analytic* functions [1, p. 83]. In this note we prove the same expansion theorem under the much weaker assumption that the coefficients σ_1, σ_2, τ_1, τ_2 are twice continuously differentiable.

2. The expansion theorem. Consequences. Our aim is to prove the following theorem.

EXPANSION THEOREM. *Let $f(z)$ be Σ-monogenic in a domain D. In the neighborhood of every point z_0 of D, $f(z)$ admits a unique (uniformly and absolutely convergent) expansion of the form* (1.9). *If $|f(z)| < M$ in D, then*

$$|a_\nu| < A^\nu M, \tag{2.1}$$

where A is a constant depending only on D, the functions σ_i, τ_j, and the distance from z_0 to the boundary of D.

With the aid of this theorem many results obtained previously for the case of analytic σ, τ can be reëstablished without the analyticity hypothesis. This applies in particular to the analogue of Morera's theorem [1, Theorem 3.6] and to the theorem on the quasi-analytic character of Σ-monogenic functions [1, p. 84]. The inequality (2.1) in conjunction with (1.10) implies the analogue of Weierstrass' convergence theorem.

The expansion theorem implies that a Σ-monogenic function possesses Σ, Σ' derivatives of *all* orders. Noting that the real part of a Σ-monogenic function is a solution of the equation

$$\{u_x\sigma_1(x)/\tau_1(y)\}_x + \{u_y\sigma_2(x)/\tau_2(y)\}_y = 0, \tag{2.2}$$

We obtain very precise information about the way in which differentiability properties of the coefficients of the elliptic equation (2.2) are reflected in the differentiability properties of solutions. We mention explicitly only an extreme case.

For

$$\sigma_1 = \sigma_2 = 1, \quad \tau_1 = \tau_2 = \exp\int\alpha(y)\,dy, \tag{2.3}$$

equation (2.2) becomes

$$\Delta u = \alpha(y)u_y. \tag{2.4}$$

On the other hand (2.3) implies that $a\cdot Z^{(n)}(z_0, x+iy_0) = a(x-x_0)^n$. Thus we obtain the following theorem: *if $\alpha(y)$ is of class C', every solution of* (2.4) *is analytic in x.*

3. Lemmas. In this section we shall prove several lemmas which are needed for the proof of the expansion theorem.

LEMMA 1. *Let D_0 be a bounded domain, D a subdomain of D_0, z_0 a point of D, r_0 the distance from z_0 to the boundary of D. Let $p(x, y)$ be a bounded continuous function defined in D_0, P a bound for $|p(x, y)|$, $\phi(x, y)$ a bounded function of class C'' defined in D and satisfying the equation.*

$$\Delta\phi = p\phi. \tag{3.1}$$

There exists a constant A depending only on D_0 and P such that the inequality

$$|\phi(x, y)| \leqq M \text{ in } D, \tag{3.2}$$

implies that

$$|\phi_x(x_0, y_0)|, |\phi_y(x_0, y_0)| \leqq AM/r_0. \tag{3.3}$$

Proof.[5] Set $\rho(x, y; \xi, \eta) = (4\pi)^{-1} \log[(x-\xi)^2 + (y-\eta)^2]$ and

$$\omega(x, y) = \int\int_D \rho(x, y; \xi, \eta) p(\xi, \eta) \phi(\xi, \eta) \, d\xi d\eta.$$

There exists a constant B depending only upon D_0 such that in D

$$|\omega(x, y)|, |\omega_x(x, y)|, |\omega_y(x, y)| \leqq BPM. \tag{3.4}$$

We have, for instance,

$$|\omega(x, y)| \leqq PM \int\int_D |\rho(x, y; \xi, \eta)| \, d\xi d\eta \leqq PM \int\int_{D_0} |\rho(x, y; \xi, \eta)| \, d\xi d\eta,$$

and the last integral has a finite maximum on the closure of D_0.

On the other hand, $h(x, y) = \phi(x, y) - \omega(x, y)$ is harmonic in D, and since, by (3.2) and (3.4), $|h| \leqq (1 + BP)M$, it follows[6] that

$$|h_x(x_0, y_0)|, |h_y(x_0, y_0)| \leqq 2(1 + BP)M/r_0. \tag{3.5}$$

From (3.4) and (3.5) the assertion follows.

Consider now a Σ-monogenic function $f = u + iv$. The function u is a solution of (2.2). Set

$$x_1 = \int_0^x [\sigma_2(t)/\sigma_1(t)]^{\frac{1}{2}} \, dt, \qquad y_1 = \int_0^y [\tau_2(t)/\tau_1(t)]^{\frac{1}{2}} \, dt,$$

and $\alpha = [\sigma_1(x)\sigma_2(x)/\tau_1(y)\tau_2(y)]^{\frac{1}{4}}$, $\phi = \alpha u$. Then ϕ considered as a function of

---

[5] The lemma is hardly new. The proof is sketched since the author knows of no reference.

[6] See [5], p. 321 where the analogous inequality is proved for three dimensions.

x_1 and y_1 satisfies the equation $\partial^2\phi/\partial x_1^2 + \partial^2\phi/\partial y_1^2 = (\partial^2\alpha/\partial x_1^2 + \partial^2\alpha/\partial y_1^2)\phi/\alpha$, so that Lemmas 1 may be applied to ϕ. A similar reasoning holds for Σ'-monogenic functions. Noting the definition of Σ-derivatives, we obtain the following result:

LEMMA 2. *Let D be a domain situated within the rectangle $a \leqq x \leqq a'$, $b \leqq y \leqq b'$, z_0 a point in D, r_0 the distance from z_0 to the boundary of D. Let $f(z)$ be Σ-monogenic (or Σ'-monogenic) in D. If in D, $|\operatorname{Re} f(z)| < M$, then $|f^{[1]}(z_0)| < AM/r_0$ where A is a constant depending only on the rectangle and on the behaviour of the functions σ, τ in the intervals $[a, a']$, $[b, b']$, respectively.*[7]

The following is an immediate consequence of Lemma 2:

LEMMA 3. *If $f_n(z)$, $n = 1, 2, \cdots$, and $f(z)$ are Σ-monogenic in a domain D, if $\operatorname{Re} f_n \to \operatorname{Re} f$ uniformly in D, and if, at some point z_0 of D, $f_n(z_0) \to f(z_0)$, then $f_n \to f$ uniformly in every closed subdomain of D.*

As another application of Lemma 2 we shall prove the inequality (2.1) of the expansion theorem for the case of a formal polynomial.

LEMMA 4. *Let D be a bounded domain containing the point z_0, $P(z)$ a formal polynomial of the form (1.6), and M an upper bound for $|P(z)|$ in D. Then*

$$|a_\nu| \leqq A^\nu M, \qquad \nu = 0, 1, \cdots, N, \tag{3.6}$$

where A is a constant of the kind described in the expansion theorem.

Proof. For $\nu = 0$ (3.6) holds for every $A \geqq 1$. Now let ν be a fixed integer $(0 < \nu \leqq N)$. Let $\bar{D}_\mu$ denote the disc $|z - z_0| \leqq (1 - \mu/\nu)r_0$ where r_0 is the distance from z_0 to the boundary of D, and $\mu = 0, 1, \cdots, \nu - 1$. Finally, let M_μ denote the maximum of $|P^{[\mu]}(z)|$ in $\bar{D}_\mu$, and let A' be the constant from Lemma 2. By repeated application of this lemma we obtain the inequalities $M_1 \leqq (\nu A'/r_0)M_0$, $M_2 \leqq (\nu A'/r_0)M_1 \leqq (\nu A'/r_0)^2 M_0, \cdots$, and finally $|P^{[\nu]}(z_0)| \leqq (A'/r_0)^\nu \nu^\nu M_0$. Since $M_0 \leqq M$ and a_ν is given by (1.7), we have $|a_\nu| \leqq (A'/r_0)^\nu (\nu^\nu/\nu!)M$, and since by Stirling's formula $\nu^\nu/\nu! \sim e^\nu(2\pi\nu)^{-\frac{1}{2}}$, the assertion follows.

The proof of the expansion theorem will be based on Lemma 4 and on the following lemma which establishes the possibility of expanding a Σ-monogenic function in a series of formal polynomials.

---

[7] An analysis of the argument will show that the maxima of the second derivatives of σ, τ enter in A.

LEMMA 5. *Let $f(z)$ be Σ-monogenic in the closed rectangle $\bar{R}$: $a \leqq x \leqq a'$ $b \leqq y \leqq b'$. Then there exists a sequence of formal polynomials $P_n(z)$ of the form* (1.6) *which converges to $f(z)$ uniformly in every closed subdomain of the open rectangle R: $a < x < a'$, $b < y < b'$.* (The point z_0 is not required to belong to $\bar{R}$).

Proof. By subtracting from $f(z)$ a formal polynomial of the fourth degree, we may obtain a function which vanishes at the vertices[8] of R. We may assume, therefore, that

$$f(a+ib) = f(a'+ib) = f(a+ib') = f(a'+ib') = 0.$$

Since the Dirichlet problem is always solvable for equation (2.2), we may represent $u = \operatorname{Re} f(z)$ in R as a sum of four solutions of (2.2), each of which vanishes on all but one side of $\bar{R}$. It will be sufficient to consider one such function. Thus we assume now that $u(x,y) = \operatorname{Re} f(z)$ satisfies the boundary conditions:

$$u(a,y) = u(a',y) = 0, \quad b \leqq y \leqq b'; \qquad u(x,b') = 0, \quad a \leqq x \leqq b',$$

$$u(x,b) = \Phi(x) \in C'', \quad a \leqq x \leqq a', \quad \Phi(a) = \Phi(a') = 0.$$

A function

$$u(x,y) = \phi(x)\psi(y) \tag{3.7}$$

will be a solution of (2.2) if

$$[\sigma_1(x)\phi'(x)]' + \lambda\sigma_2(x)\phi(x) = 0, \tag{3.8}$$

$$[\psi'(y)/\tau_2(y)]' - \lambda\psi(y)/\tau_1(y) = 0. \tag{3.9}$$

The characteristic values of (3.8) for the boundary conditions $\phi(a) = \phi(a') = 0$ are simple and positive. Let us denote them by Λ_j^2, $j = 1, 2, \cdots$, and let $\phi_j(x)$ be the corresponding (normalized) characteristic functions. Since Λ_j^2 is certainly not a characteristic value of (3.9) for the boundary conditions $\psi(b) = \psi(b') = 0$, there exists a solution $\psi_j(y)$ of (3.9) for $\lambda = \Lambda_j^2$ which satisfies the conditions $\psi_j(b) = 1$, $\psi_j(b') = 0$. It is well known that $\Phi(x)$ admits the uniformly and absolutely convergent expansion $\Phi(x) = \sum_{j=1}^{\infty} \alpha_j\phi_j(x)$, $a \leqq x \leqq a'$. The series $\sum_{j=1}^{\infty} \alpha_j\phi_j(x)\psi_j(y)$ converges to $u(x,y)$ uniformly on the boundary of R. The same must be true throughout R since solutions of (2.2) are known to obey the maximum-minimum principle.

---

[8] See the interpolation theorem in [1], p. 89. This theorem is proved without using analyticity of the functions σ, τ.

Now we determine Σ-monogenic functions $f_j(z)$ such that

$$\text{(3.10)} \qquad \operatorname{Re} f_j(z) = \alpha_j \phi_j(x)\psi_j(y), \qquad j = 1, 2, \cdots,$$

$$\text{(3.11)} \qquad \operatorname{Im} f_1(z_1) = \operatorname{Re} f(z_1), \quad \operatorname{Im} f_j(z_1) = 0, \qquad j = 2, 3, \cdots,$$

where z_1 is some fixed point in R. By Lemma 3 we conclude that $f(z) = \sum_{j=1}^{\infty} f_j(z)$, the series being uniformly convergent in every closed subdomain of R.

The assertion of Lemma 5 would be proved if we could find for every positive ϵ and for every j a formal polynomial of the form (1.6) which in R differs from $f_j(z)$ by less than ϵ. That this can be done, however, follows from the fact that every Σ-monogenic function satisfying (3.10) can be expanded in a formal power series of the form (1.9) which converges uniformly in every bounded domain.[9]

4. Proof of the expansion theorem.[10] Let $f(z)$ be Σ-monogenic in the domain D containing z_0, and let $|f(z)|$ not exceed M in D. According to Lemma 5 there exists a sequence of formal polynomials,

$$P_n(z) = \sum_{\nu=0}^{N_n} a_\nu^{(n)} \cdot Z^{(\nu)}(z_0; z),$$

such that

$$\text{(4.1)} \qquad f(z) = \lim_{n\to\infty} P_n(z),$$

uniformly in the closure of some neighborhood G_0 of z_0. We set $a_\nu^{(n)} = 0$ for $\nu > N_n$. Let A' denote the constant of Lemma 4 for the domain G_0 and the point z_0. Then, by (4.1) and Lemma 4,

$$\text{(4.2)} \qquad |a_\nu^{(n)}| \leqq A'^{\nu}(M + \epsilon_n); \qquad \lim_{n\to\infty} \epsilon_n = 0.$$

The limit

$$\text{(4.3)} \qquad \lim_{n\to\infty} a_\nu^{(n)} = a_\nu$$

exists for $\nu = 0, 1, \cdots$. In fact, we have that

$$\text{(4.4)} \qquad |a_\nu^{(n)} - a_\nu^{(m)}| \leqq A'^{\nu} \cdot \max |P_n(z) - P_m(z)| \,(z \in \bar{G}_0).$$

From (4.2) we conclude that

$$\text{(4.5)} \qquad |a_\nu| \leqq A'^{\nu} M,$$

---

[9] See [1], Theorem 10.2. The proof of this theorem uses no analyticity assumptions.

[10] I am indebted to Professor Hermann Weyl for suggesting a simplification of the argument.

so that the formal power series

$$\sum_{\nu=0}^{\infty} a_\nu \cdot Z^{(\nu)}(z_0; z) \tag{4.6}$$

converges uniformly and absolutely in some neighborhood G_1 of z_0. Let r be a positive number such that the disc $G: |z - z_0| < r$ is contained in G_1 and $rA'C(2|z_0| + r) < 1$, where C is the function in inequality (1.8). For $m \to \infty$, (4.4) yields

$$|a_\nu - a_\nu^{(n)}| \leqq A'^\nu \cdot \max |P_n(z) - f(z)| = A'^\nu \eta_n, \qquad \eta_n \to 0.$$

Using this inequality, we see from (1.5) and (1.8) that, for z in G,

$$\left|\sum_{\nu=0}^{\infty} a_\nu \cdot Z^{(\nu)}(z_0; z) - P_n(z)\right| \leqq \sum_{\nu=0}^{\infty} |(a_\nu - a_\nu^{(n)}) \cdot Z^{(\nu)}(z_0; z)|$$

$\leqq \eta_n/\{1 - rA'C(2|z_0| + r)\}$, so that, by (4.1), $f(z) = \sum_{\nu=0}^{\infty} a_\nu \cdot Z^{(\nu)}(z_0; z)$. The uniqueness of this expansion follows at once from Lemma 4.

It remains to prove inequality (2.1). (Note that it is not implied by (4.5) since A' depends upon G_0). We observe first that since $f(z)$ may be expanded in a formal power series in the neighborhood of every point of D, all Σ, Σ' derivatives of f exist. Next we repeat verbatim the proof of Lemma 4, replacing P by f and using (1.10) instead of (1.7).

The proof of the expansion theorem is now complete. The method used to establish this result ought to suffice to free the expansion theorems obtained by Diaz [4] and Protter [6] for their generalizations of Σ-monogenic functions from unnecessary analyticity conditions. It remains an open question, however, whether the conditions imposed by us on the functions σ, τ can be substantially weakened.

SYRACUSE UNIVERSITY.

REFERENCES.

[1] L. Bers and A. Gelbart, "On a class of functions defined by partial differential equations," *Transactions of the American Mathematical Society*, vol. 56 (1944), pp. 67-93.

[2] ———, "On a class of differential equations in mechanics of continua," *Quarterly of Applied Mathematics*, vol. 1 (1943), pp. 168-188.

[3] ———, "On generalized Laplace transformations," *Annals of Mathematics*, vol. 48 (1947), pp. 342-357.

[4] J. B. Diaz, "On a class of partial differential equations of even order," *American Journal of Mathematics*, vol. 68 (1946), pp. 611-659.

[5] O. D. Kellogg, *Foundations of Potential Theory*, Berlin, 1929.

[6] M. H. Protter, "Generalized spherical harmonics," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 314-341.

ON THE SPECTRUM OF A BOUNDARY VALUE PROBLEM WITH TWO SINGULAR ENDPOINTS.*

By KENNETH G. WOLFSON.

1. Let a positive $p(t)$ and a real-valued $q(t)$, where $-\infty < t < \infty$, be continuous functions with the property that the differential equation

$$(px')' + (q + \lambda)x = 0 \tag{1}$$

is of the "Grenzpunkt" type for both $t = -\infty$ and $t = +\infty$. Thus (1) determines a boundary value problem in the Hilbert space $L^2(-\infty, +\infty)$; that is, $L(z) = -(pz')' - qz$ is a self-adjoint operator defined on the set of functions $z(t)$ for which z, pz' are absolutely continuous; and z, $L(z)$ are of class $L^2(-\infty, \infty)$. Let the spectrum of this operator be denoted by S and the set of its cluster points by S'. The assumptions above also imply that (1) and a fixed homogeneous boundary condition,

$$x(c)\cos\alpha + x'(c)\sin\alpha = 0, \tag{2}$$

where $-\infty < c < \infty$ and $0 \leqq \alpha < \pi$, determine two boundary value problems, one in each of the Hilbert spaces $L^2(-\infty, c)$ and $L^2(c, \infty)$. Let $S^{c\alpha}, S_{c\alpha}$ denote the spectra of the respective problems. The derived sets $S^{c\alpha\prime}$, $S_{c\alpha}{}'$ are independent of c and α ([8], p. 251) and may be denoted by S^*, S_*, respectively.

The object of this paper will be to determine the relationship between the sets S', S^*, S_* and those between the set S and the spectra $S_a{}^b$ of the Sturm-Liouville problems belonging to (1) and the boundary conditions

$$x(a) = 0 \quad \text{and} \quad x(b) = 0, \tag{3}$$

where $-\infty < a < b < \infty$.

The second part of the problem is analogous to the corresponding problem concerning $S_{c\alpha}$ treated in [2]. The results below will be dependent on those of [2].

For $\lambda_1 < \lambda_2$ and $a < b$, let $N(a, b, \lambda_1, \lambda_2)$ be the number of points of $S_a{}^b$ in the closed interval $[\lambda_1, \lambda_2]$. Put $n(\lambda_1, \lambda_2) = \liminf N(a, b, \lambda_1, \lambda_2)$, where $a \to -\infty$ and $b \to +\infty$. The following theorems will be proved:

* Received March 29, 1950. Presented to the American Mathematical Society, April 28, 1950.

(I) *A necessary and sufficient condition for* $\lambda = \lambda_0$ *to be in* S *is that*

$$n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1 \text{ for every } \epsilon > 0. \tag{4}$$

(II) *A necessary and sufficient condition for* $\lambda = \lambda_0$ *to be in* S' *is that*

$$n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) = \infty \text{ for every } \epsilon > 0. \tag{5}$$

Let $\delta(\lambda_0)$ denote the number (0, 1 or 2) of the half lines $(-\infty, 0]$, $[0, \infty)$ on which (1) is oscillatory for λ near, but exceeding, λ_0. Put $m(\lambda_1, \lambda_2) = \limsup N(a, b, \lambda_1, \lambda_2)$, where $a \to -\infty$ and $b \to \infty$. It turns out that condition (4) is equivalent to

$$m(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq \delta(\lambda_0) + 1 \text{ for every } \epsilon > 0. \tag{4'}$$

It may be pointed out that, in (I), the boundary conditions (3) cannot be replaced by arbitrary homogeneous boundary conditions, say

$$x(a)\cos\beta + x'(a)\sin\beta = 0, \qquad x(b)\cos\beta + x'(b)\sin\beta = 0. \tag{6}$$

Corresponding to the differential equation $x'' + \lambda x = 0$ and the boundary conditions (6) determined by $\beta = \pi/4$, the number $\lambda = -1$ is an eigenvalue for every choice of a, b; but the set S is the half-line $0 \leqq \lambda < \infty$, and does not contain the point $\lambda = -1$.

It will be clear that the proof of (I) can be modified so as to imply the following theorem. This theorem is the analogue of one concerning S_{ca}, which was proved first in [8], pp. 252-257, with restrictions on $q(t)$ and λ_2, and without restrictions in [2], p. 917.

(III) *If* $n(\lambda_1, \lambda_2) < \infty$, *put* $N(\lambda) = n(\lambda_1, \lambda)$ *for* $\lambda_1 < \lambda < \lambda_2$. *Then* (i) $N(\lambda)$ *is a non-decreasing (integral valued) step-function*; (ii) $N(\lambda)$ *is continuous from the left*; (iii) *every jump (if any) of* $N(\lambda)$ *has the value* 1; *finally,* (iv) *the discontinuity points of* $N(\lambda)$ *on* $\lambda_1 < \lambda < \lambda_2$ *are in the point spectrum and no other* λ *values of the open interval* (λ_1, λ_2) *are in the spectrum* S *of* (1).

The following theorem will also be proved:

(IV) $S' = S_* + S^*$, *where the addition is set-theoretic.*

The set S' can be decomposed into the cluster points of the point spectrum and the continuous spectrum C. According to Weyl ([9], p. 450), each λ in C is at most of double multiplicity, as defined by Hellinger [5]. If $p \equiv 1$, while q is even and periodic, then it is known [10] that the sets S_* and S^* are identical and consist entirely of continuous spectra. Furthermore every

point λ of $S_* = S^*$ is also in C with multiplicity two; cf. the spectral resolution given in [6], p. 934. Thus suggests the possibility that, in the general case, (IV) has some interpretation as regards the multiplicities of points of the continuous spectrum C. This will remain an open question.

As a consequence of (IV), results on the set S^* can be carried over to the set S'. For example, by the use of the theorem (*) of [3] we have:

If there exists a number μ such that for $\lambda < \mu$ every solution of (1) has only a finite number of zeros for $-\infty < t < \infty$, but has infinitely many zeros when $\lambda > \mu$, then μ is the least point of S'. If such a number μ fails to exist, then S clusters at $\lambda = -\infty$.

2. Proof of the necessity of (4) in (I). If $\liminf N(a, b, \lambda_1, \lambda_2) = k$, where $a \to -\infty$ and $b \to +\infty$, then the open interval (λ_1, λ_2) contains at most k points of S.

This can be established as follows: Let K_{ab} denote the bounded integral operator whose kernel on the rectangle $a \leqq s,\ t \leqq b$ is the Green function determined by (1) for $\lambda = i$ and boundary conditions (3). This operator determines the spectrum S_a^b. Let K denote the corresponding integral operator for the interval $(-\infty, \infty)$; so that $K = (L - iI)^{-1}$, where I is the identity operator. For convenience, suppose that the domain of K_{ab} is not $L^2(a, b)$ but $L^2(-\infty, +\infty)$; and that, for any element g of this Hilbert space, $K_{ab}g$ is 0 outside the interval $[a, b]$. From the form of the Green functions and from the definition of "Grenzpunkt" ([8], p. 226), it can be verified that K_{ab} converges to K, as $a \to -\infty$ and $b \to +\infty$, in the sense of strong convergence of bounded operators on the Hilbert space $L^2(-\infty, \infty)$; cf. the derivation of the spectral resolution of L as given by Titchmarsh [7], pp. 56-57. The statement at the beginning of this section follows from known properties of strongly convergent bounded operators.

This assures the necessity of (4).

3. Proof of sufficiency of (4) in (I). The proof will be divided into three parts (i), (ii), (iii) corresponding to the cases $\delta(\lambda_0) = 0, 1, 2$ respectively.

(i) If $n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1$ for a given $\epsilon > 0$, then there exist numbers A_ϵ, B_ϵ such that, for all $a \leqq A_\epsilon$ and $b \geqq B_\epsilon$, the set S_a^b has at least one point in the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. By Theorem (I) in [2], the set S^b, where $S^b = S^{b0}$, contains at least one point in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ for every $b \geqq B_\epsilon$. By the Lemma in 3 of [1], there exist solutions of (1) of class $L^2(-\infty, b)$, corresponding to the values λ lying in the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ for

sufficiently small $\epsilon > 0$. Since (1) is of the *Grenzpunkt* type at $t = -\infty$, these solutions are unique up to constant factors. Denote by $x(t, \lambda)$ a solution of (1) which is $L^2(-\infty, b)$. Since $\delta(\lambda_0) = 0$, the solution $x(t, \lambda)$ has only a finite number of zeros on $-\infty < t < \infty$. The proof can be completed by adapting the method used in [1] for the analogous problem concerning the sets $S_{ca}{}^b$ and S_{ca}, for fixed c. That method is based on the following: If $x^*(t, \lambda)$ denotes a solution of (1) and (2), then the n-th largest zero of x^* on $c < t < \infty$ is a continuous decreasing function of λ, and the Wronskian of $x^*(t, \lambda^1)$ and $x^*(t, \lambda^2)$ is zero at $t = c$. In the present case of the function $x(t, \lambda)$, the point $t = -\infty$ replaces the point $t = c$. The n-th largest zero of $x(t, \lambda)$ on $-\infty < t < \infty$ is a continuous decreasing function of λ. Also, the Wronskian of $x(t, \lambda^1)$ and $x(t, \lambda^2)$ satisfies

$$(7) \qquad p(t)[x'(t, \lambda^2)x(t, \lambda^1) - x(t, \lambda^2)x'(t, \lambda^1)] \to 0 \text{ as } t \to -\infty;$$

[8], p. 242; cf. formula (19) in [4].

The details of the proof for case (i) will not be given as they follow closely those of [1].

(ii) In the case $\delta(\lambda_0) = 1$, it can be assumed that (1) is non-oscillatory on the half axis $-\infty < t \leqq 0$. Suppose, if possible, that (4) holds but that λ_0 is not in S. Then there exists an $\epsilon > 0$ such that the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ is free of points of S. Since $n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1$, there exist numbers A_ϵ, B_ϵ with the property that, whenever $a \leqq A_\epsilon$, $b \geqq B_\epsilon$, the set $S_a{}^b$ has at least one point in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. Theorem (I) in [2] implies that, for every $b \geqq B_\epsilon$, the set S^b has at least one point in the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. It will be shown that there exist values of b corresponding to which S^b has at least two points in that interval. Let $\{\mu_h(b)\}$ denote the set of eigenvalues in S^b which are less than the least point of $S^{b'} = S^*$. The continuity and monotony of the n-th zero of $x(t, \lambda)$, mentioned above, and the fact that the number of zeros on $(-\infty, b)$ of eigenfunctions corresponding to successive elements of the sequence $\{\mu_h(b)\}$ differ by one, allow an adaptation of the arguments of [2], p. 916, as follows:

Let $x(t, \lambda)$ have the same meaning as in (i). Consider a value $b' > b$ satisfying $x(b', \lambda_0 - \epsilon) = 0$; so that $\lambda_0 - \epsilon = \mu_h(b')$ for some value of h. Then $\mu_{h+1}(b') \leqq \lambda_0 + \epsilon$. For if this were not the case, it would follow that $\mu_h(b'') < \lambda_0 - \epsilon$, while $\mu_{h+1}(b'') > \lambda_0 + \epsilon$ for b'' near, but exceeding, b'. This contradicts the fact that for every $b \geqq B_\epsilon$ the set S^b contains at least one element in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. Hence, for some arbitrarily large b, S^b has at least two points in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$.

To complete the proof, it will be shown that the assumption that no point of S is in the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ leads to a contradiction.

Since the addition of a constant λ_0 to $q(t)$ simultaneously translates the sets S^b, S by $-\lambda_0$, it can be assumed that $\lambda_0 = 0$. Thus, for some choice of b and h, the numbers $\mu_h(b)$, $\mu_{h+1}(b)$ lie in the interval $[-\epsilon, \epsilon]$, but this interval contains no points of S. Let $x_h(t)$, $x_{h+1}(t)$ be eigenfunctions corresponding to μ_h, μ_{h+1} respectively. Put

$$z(t) = c_1 x_h(t) + c_2 x_{h+1}(t) \text{ or } z(t) = 0 \tag{8}$$

according as $-\infty < t \leqq b$ or $b < t < \infty$, where c_1, c_2 are chosen so that $z'(b-0) = 0$ (and $z(t) \not\equiv 0$). Then z, pz' are absolutely continuous, and z, $L(z)$ are of class $L^2(-\infty, \infty)$. Hence $z(t)$ belongs to the domain of the self-adjoint operator $L(z)$. From (1) and (8)

$$L(z) = c_1 \mu_h x_h(t) + c_2 \mu_{h+1} x_{h+1}(t) \text{ or } L(z) = 0,$$

according as $-\infty < t \leqq b$ or $b < t < \infty$. Since $|\mu_h|, |\mu_{h+1}| \leqq \epsilon$, it follows from (8) that

$$\int_{-\infty}^{+\infty} L^2(z)\,dt \leqq \epsilon^2 \int_{-\infty}^{+\infty} z^2 dt, \tag{9}$$

since x_h, x_{h+1} are orthogonal on $(-\infty, b)$. Since $[-\epsilon, \epsilon]$ is free of points of S, the spectral resolution of the operator $L(z)$ implies

$$\int_{-\infty}^{+\infty} L^2(z)\,dt > \epsilon^2 \int_{-\infty}^{+\infty} z^2 dt. \tag{10}$$

This inequality contradicts (9) and completes the proof of the case (ii).

(iii) Assume that $\delta(\lambda_0) = 2$, that (4) holds and that, if possible, λ_0 is not in S. Then there is an interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ free of points of S; and there exist numbers A_ϵ, B_ϵ such that for all $a \leqq A_\epsilon$ and $b \geqq B_\epsilon$, the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ contains at least one point of $S_a{}^b$. The theorem (I) in [2] implies that $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ contains at least one point $\lambda = \lambda_1$ of S^b, and at least one point $\lambda = \lambda_2$ of S_a. Let $x_1 = x_1(t)$, $x_2 = x_2(t)$ denote corresponding eigenfunctions. Let $d < a$ be a zero of x_1 and $e > b$ be a zero of x_2. These zeros exist in view of the assumption $\delta(\lambda_0) = 2$. Let $x_3 = x_3(t)$ denote an eigenfunction corresponding to a point $\lambda = \lambda_3$ of $S_d{}^e$ lying in the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. Finally, let $z(t) = c_1 x_1(t), x_3(t), c_2 x_2(t)$ for $-\infty < t \leqq d$, $d < t \leqq e$, and $e < t < \infty$, respectively, where c_1, c_2 are chosen so as to make $z'(t)$ continuous. Since x_1, x_2, x_3 satisfy (1) for $\lambda = \lambda_1, \lambda_2, \lambda_3$, respectively, it follows that $(pz')' = 0$ for $t = d, e$. Hence $(pz')'$ is continuous; so that z, pz' are absolutely continuous on $-\infty < t < \infty$. Thus

$z(t)$ is in the domain of the operator $L(z)$. By a procedure analogous to that used in (ii), one obtains a contradiction to the fact that $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ is free of points of S. This completes the proof of case (iii).

4. Proof of (II). The necessity of (5) follows from 2. The proof of (I) can be modified so as to imply that if $n(\lambda_1, \lambda_2) = k$, then there are k points of S in $[\lambda_1, \lambda_2]$; cf. (III). This shows the sufficiency of (5).

5. Proof of (IV). Assume that λ_0 is in S_*. Theorem (I) of [2] implies that if c is fixed, $-\infty < c < \infty$, then there exist corresponding sets $S_c{}^b$ having arbitrarily many points in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ for every $\epsilon > 0$. Hence $n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) = \infty$ for every $\epsilon > 0$, and so λ_0 is in S' by (II). Similarly, if λ_0 is in S^*, it is in S'.

Assume that λ_0 is in S'. It remains to verify that λ_0 is in either S^* or S_*. Suppose, if possible, that this is not the case. If c is a fixed number, there exists an interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ which contains at most one point of each of the sets S^c, S_c. Then no set $S_a{}^c$ or $S_c{}^b$ for any $a < c < b$ can have more than two points in the interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. But then Sturm's comparison theorem shows that $N(a, b, \lambda_0 - \epsilon, \lambda_0 + \epsilon) \leqq 5$ for all $a < c < b$, which is contradictory to the fact that λ_0 is assumed to be in S', and so $n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) = \infty$.

6. Proof of equivalence of conditions (4) and (4'). The proof will again be divided into three parts (i), (ii), (iii) corresponding to $\delta(\lambda_0) = 0, 1, 2$ respectively.

(i) In this case it is only necessary to show that if, for some $\epsilon > 0$, the limit superior $m(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1$, then $n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1$, for the converse is obvious. The assumption $m(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1$ implies that there exists a sequence of intervals $(a_1, b_1), (a_2, b_2), \cdots$ such that $a_k > a_{k+1} \to -\infty$ and $b_k < b_{k+1} \to \infty$, as $k \to \infty$, and that $S_a{}^b$, where $(a, b) = (a_k, b_k)$, has at least one point $\lambda = \lambda_k$ in the closed interval $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$. Let p_k denote the number of zeros in (a_k, b_k) of the eigenfunction corresponding to $\lambda = \lambda_k$. Since $\delta(\lambda_0) = 0$, the sequence $p_1, p_2, \cdots$ is bounded. Hence it can be supposed that $p_k = p$ is independent of k; otherwise the sequence (a_1, b_1), $(a_2, b_2), \cdots$ is replaced by a subsequence. If $\mu_p(a, b)$ denotes that point of $S_a{}^b$, which corresponds to an eigenfunction having exactly p zeros on (a, b), Sturm's comparison theorem shows that $\mu_p(a', b') < \mu_p(a, b)$ whenever $a' < a$ and $b' > b$. Since $\lambda_k = \mu_p(a_k, b_k)$, it follows that $\lambda_1 > \lambda_2 > \cdots > \lambda_0 - \epsilon$. The assertion to be proved is an obvious consequence.

(ii) The proof used in (ii) of **3** shows that if $n(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 1$, then $m(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 2$; cf. [2], p. 916. The converse statement follows from Sturm's comparison theorem.

(iii) The equivalence of (4) and (4′) in this case can be shown by proving that $m(\lambda_0 - \epsilon, \lambda_0 + \epsilon) \geqq 3$ is necessary and sufficient for λ_0 to be in S. The sufficiency follows from an obvious modification of the construction used in part (iii) of **3**. To show the necessity, assume that λ_0 is in S. Then it belongs to both S^c and S_c, where $t = c$ is a zero of an eigenfunction corresponding to λ_0. Thus, for any $\epsilon > 0$, there exist arbitrarily large numbers $-a, b$ such that $S_a{}^c$ and $S_c{}^b$ each have at least two points in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$; cf. [2], p. 918. Then $S_a{}^b$ has at least three points in $[\lambda_0 - \epsilon, \lambda_0 + \epsilon]$ by Sturm's comparison theorem. This completes the proof.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman, "Differential equations with non-oscillatory eigenfunctions," *Duke Mathematical Journal*, vol. 15 (1948), pp. 697-709.

[2] ———, "A characterization of the spectra of one dimensional wave equations," *American Journal of Mathematics*, vol. 71 (1949), pp. 915-920.

[3] ——— and C. R. Putnam, "The least cluster point of the spectrum of boundary value problems," *ibid.*, vol. 70 (1948), pp. 849-855.

[4] ——— and A. Wintner, "A separation theorem for continuous spectra," *ibid.*, vol. 71 (1949), pp. 650-662.

[5] E. Hellinger, *Die Orthogonalinvarianten quadratischer Formen von unendlichen Variablen* (Inaugural-Dissertation), Göttingen 1907.

[6] K. Kodaira, "The eigenvalue problem for ordinary differential equations of the second order and Heisenberg's theory of S-matrices," *American Journal of Mathematics*, vol. 71 (1949), pp. 921-945.

[7] E. C. Titchmarsh, *Eigenfunction expansions associated with second order differential equations*, Oxford (1946).

[8] H. Weyl, "Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen," *Mathematische Annalen*, vol. 68 (1910), pp. 220-269.

[9] ———, "Ueber gewöhnliche Differentialgleichungen mit singulären Stellen und ihre Eigenfunktionen," *Göttinger Nachrichten* (1910), pp. 442-467.

[10] A. Wintner, "Stability and spectrum in the wave mechanics of lattices," *Physical Review*, vol. 72 (1947), pp. 81-82.

ON CUBE-COVERINGS OF n-SPACE.*

By L. G. PECK.

It was shown by Hajos[1] in 1941 that in every covering of Euclidean n-space by unit cubes whose edges are parallel to the coordinate axes and whose centers form a lattice (modul) there is a pair of cubes with a complete $(n-1)$-dimensional face in common. The present note contains the results of an attempt to generalize the above theorem by omitting the condition that the centers of the cubes form a lattice.

$\mathfrak{R}_n$ will denote the n-dimensional vector space over the real field; small German letters will be used for elements of $\mathfrak{R}_n$, and the corresponding Latin letters with subscripts for their components (*e. g.*, $\mathfrak{x} = (x_1, \cdots, x_n)$, $\mathfrak{a}_1 = (a_{11}, \cdots, a_{1n})$). The unit vectors, $\mathfrak{e}_1, \cdots, \mathfrak{e}_n$, are defined by the equation $e_{ij} = 0$ or 1, according as $i \neq j$ or $i = j$. There will also be occasion to use the set Δ_n of n-tuples $\delta = (\delta_1, \cdots, \delta_n)$, where $\delta_i = \pm 1 (i = 1, \cdots, n)$. Note that Δ_n contains 2^n elements.

The following definition is introduced to make precise the notion of a cube-covering of n-space: *A subset $\mathfrak{A}$ of $\mathfrak{R}_n$ will be called an n-grill if, for every $\mathfrak{x} \in \mathfrak{R}_n$, there exists a unique $\mathfrak{a} \in \mathfrak{A}$ such that* $0 \leqq x_i - a_i < 1$, $(i = 1, \cdots, n)$.

It is clear how one can associate a cube-covering with every n-grill, and an n-grill with every cube-covering. Notice that the defining property of a grill is unaltered by translation, so that there is no loss of generality in assuming that a particular point of the grill is the origin $\mathfrak{o} = (0, \cdots, 0)$.

A property (P) of grills, which will be used repeatedly in the sequel, is as follows:

If $\mathfrak{A}$ is an n-grill, if $\mathfrak{a}, \mathfrak{a}' \in \mathfrak{A}$, and if $|a_i - a_i'| < 1$, $(i = 1, \cdots, n)$, *then* $\mathfrak{a} = \mathfrak{a}'$.

Proof. Let $x_i = \max(a_i, a_i')$ for $i = 1, \cdots, n$, and apply the definition of $\mathfrak{A}$.

* Received February 9, 1950.

[1] Georg Hajos, "Über einfache und mehrfache Bedeckung des n-dimensionalen Raumes mit einem Würfelgitter," *Mathematische Zeitschrift*, vol. 47 (1941). The reader should also consult papers by Keller, *Journal für die Reine und Angewandte Mathematik*, vol. 163 (1930), and Perron, *Mathematische Zeitschrift*, vol. 46 (1940).

It is now possible to prove a lemma, which, interpreted geometrically, states that the fundamental half-open cube which is used in the covering may be constructed from a closed cube by omitting from it any set of 2^{n-1} non-parallel $(n-1)$-dimensional faces. (The difficulty of making this last statement clearer without departing from geometrical terminology accounts for the introduction of the concept of a grill.)

LEMMA. *Let* $\mathfrak{A}$ *be an* n*-grill,* $\mathfrak{x} \in \mathfrak{R}_n$, *and* $\delta \in \Delta_n$. *Then there exists a unique* $\mathfrak{a} \in \mathfrak{A}$ *such that* $0 \leqq x_i + \delta_i a_i < 1$; $(i = 1, \cdots, n)$.

Proof. The uniqueness of $\mathfrak{a}$ follows from (P). For the existence, we can find (by definition) elements, $\mathfrak{a}_1, \mathfrak{a}_2, \mathfrak{a}_3, \cdots$, of $\mathfrak{A}$ which satisfy

$$\begin{aligned} 0 \leqq x_i - a_{ki} < 1 \qquad & \text{if } \delta_i = -1 \\ 0 \leqq (-x_i - 1 - 1/k) - a_{ki} < 1 \qquad & \text{if } \delta_i = +1 \end{aligned} \qquad (k = 1, 2, \cdots).$$

Since $| a_{ki} - a_{li} | < 2$ $(i = 1, \cdots, n;\ k, l = 1, 2, \cdots)$, it follows from (P) that there is only a finite number of distinct elements in the sequence $\mathfrak{a}_1, \mathfrak{a}_2, \cdots$. Also, if $\mathfrak{a}_k = \mathfrak{a}_l = \mathfrak{b}$, and $k < l$, we have $0 \leqq x_i - b_i < 1$ if $\delta_i = -1$, and $1/k \leqq (-x_i - 1) - b_i < 1 + (1/l)$ if $\delta_i = +1$. Therefore, if $k < m < l$, we have $| b_i - a_{mi} | < 1$, $(i = 1, \cdots, n)$, and $\mathfrak{a}_m = \mathfrak{b}$ by virtue of (P). Thus, for a suitably large K, $\mathfrak{a}_K = \mathfrak{a}_{K+1} = \cdots = \mathfrak{a}$, say. This $\mathfrak{a}$ clearly satisfies the conditions of the lemma.

With the above preparation, it will now be shown that in any cube covering of the type under consideration each cube has a vertex in common with some other cube. This statement, it should be remarked, is no longer true if the word "vertex" is replaced by the word "edge," trivial counter-examples being available in three dimensions.[2]

THEOREM. *Let* $\mathfrak{A}$ *be an* n*-grill and* $\mathfrak{p} \in \mathfrak{A}$. *Then there exist numbers,* $\epsilon_1, \cdots, \epsilon_n$, *not all zero, such that* $\epsilon_i = 0$ *or* 1 $(i = 1, \cdots, n)$ *and* $\mathfrak{p} + \epsilon_1 \mathfrak{e}_1 + \cdots + \epsilon_n \mathfrak{e}_n \in \mathfrak{A}$.

Proof. Without loss of generality, take $\mathfrak{p} = \mathfrak{o}$. Denote by $\mathfrak{A}_0$ the set of $\mathfrak{a} \in \mathfrak{A}$ which satisfy $0 \leqq a_i \leqq 1$ $(i = 1, \cdots, n)$. If $\delta \in \Delta_n$, the preceding lemma assures the unique existence of $\mathfrak{a}(\delta) \in \mathfrak{A}$ satisfying $0 \leqq (1 - \delta_i)/2 + \delta_i a_i(\delta) < 1$; $(i = 1, \cdots, n)$, *i. e.*, $0 \leqq a_i(\delta) < 1$ when $\delta_i = +1$, and

[2] Consider the 3-grill whose elements are triples of integers (x, y, z) with the following exceptions: when $x = 0$, $y = \pm 1$, take the third component of the form $z + \frac{1}{2}$, z integral; when $y = 0$, $z = \pm 1$, take the first component of the form $x + \frac{1}{2}$, x integral; when $z = 0$, $x = \pm 1$, take the second component of the form $y + \frac{1}{2}$, y integral. $(0, 0, 0)$ is an element of this grill, but the associated cube has no edge in common with any other cube of the covering.

$0 < a_i(\delta) \leqq 1$ when $\delta_i = -1$. The function $\mathfrak{a}(\delta)$ thus maps each element of Δ_n onto an elements of $\mathfrak{A}_0$. Consider the inverse mapping. If $\mathfrak{b} \in \mathfrak{A}_0$, we will have $\mathfrak{b} = \mathfrak{a}(\delta)$ on condition that $0 \leqq b_i < 1$ when $\delta_i = +1$, and $0 < b_i \leqq 1$ when $\delta_i = -1$.

In other words, $\delta_i = +1$ when $b_i = 0$, $\delta_i = -1$ when $b_i = 1$, and finally $\delta_i = \pm 1$ when $0 < b_i < 1$. Consequently, *if* $\zeta(\mathfrak{b})$ *denotes the number of components,* b_i, *of* $\mathfrak{b}$ *which satisfy* $0 < b_i < 1$, then the inverse of the mapping $\mathfrak{a}(\delta)$ maps each element $\mathfrak{b} \in \mathfrak{A}_0$ onto $2^{\zeta(\mathfrak{b})}$ elements of Δ_n. Thus, $\Sigma 2^{\zeta(\mathfrak{b})} = 2^n$, where the summation runs over every $\mathfrak{b} \in \mathfrak{A}_0$. Since $\zeta(\mathfrak{o}) = 0$, this means that $\Sigma' 2^{\zeta(\mathfrak{b})} = 2^n - 1$ where the $'$ refers to the omission of $\mathfrak{b} = \mathfrak{o}$. Consequently $\Sigma' 2^{\zeta(\mathfrak{b})}$ is odd and so $\zeta(\mathfrak{b}) = 0$ for some $\mathfrak{b} \in \mathfrak{A}_0$, $\mathfrak{b} \neq \mathfrak{o}$. This completes the proof.

THE JOHNS HOPKINS UNIVERSITY.

ON THE LOCAL UNIQUENESS OF GEODESICS.*

By PHILIP HARTMAN.

1. Statement of the theorems. Let $E(u, v)$, $F(u, v)$, $G(u, v)$ be defined in a neighborhood of $(u, v) = (0, 0)$ in such a way that

$$ds^2 = Edu^2 + 2Fdudv + Gdv^2 \tag{1}$$

is a positive-definite element of arc-length. If E, F, G are of class C^1, then the Christoffel symbols Γ^i_{jk} exist and are continuous functions of (u, v) and so the differential equations

$$u''_i + \Gamma^i_{jk} u'_j u'_k = 0 \text{ for } i = 1, 2, \qquad (u_1, u_2) = (u, v) \tag{2}$$

are meaningful and define the geodesics associated with (1). Professor Wintner has raised the question of whether or not geodesics are uniquely determined (locally) by (2) and by a set of initial conditions (even though the coefficient functions Γ^i_{jk} are not subject to a "uniqueness" restriction, say of the Lipschitz type). It turns out that the answer to this question is of a very surprising nature.

On the one hand, if it is only assumed, as above, that E, F, G are of class C^1, then the solutions of (2) are not in general uniquely determined by initial conditions. On the other hand, if (1) is the first fundamental form of a (2-dimensional) surface of class C^2 in a 3-dimensional Euclidean space, then the solutions of (2) are unique. This last statement can be reworded in the following manner:

THEOREM 1. *On a surface of class C^2, every geodesic is uniquely determined (locally) by its initial conditions.*

The proof of Theorem 1 will depend on the circumstance, pointed out by Weyl [5], pp. 42-44 (cf. also van Kampen [4], p. 135), that the theorem of Gauss-Bonnet and/or the Theorema Egregium (in a suitable form) holds on surfaces of class C^2.

Let X denote the vector (x, y, z) and let $X = X(u, v)$ be of class C^n, where $n \geqq 1$, on a (u, v) domain and such that the vector product $(X_u, X_v) \neq 0$. Then $X = X(u, v)$ is called a surface of class C^n.

The proof of Theorem 1 can be refined to show that it is possible to introduce geodesic polar coordinates (r, ϕ) at any point on a surface of class

* Received December 28, 1949.

C^2 and that the resulting transformation $(r, \phi) \to (u, v)$ is of class C^1. In the standard textbooks on differential geometry, it is assumed that the surface $X = X(u, v)$ is of class C^3 in order to introduce geodesic polar coordinates. Their result, under that assumption, is merely that the transformation $(r, \phi) \to (u, v)$ is of class C^1. The proof below will show that if the surface is of class C^n, where $n \geqq 2$, then the mapping $(r, \phi) \to (u, v)$ is of class C^{n-1}.

THEOREM 2. *To every point P on a surface $X = X(u, v)$ of class C^n, where $n \geqq 2$, there corresponds a neighborhood N of P with the property that, for any point $Q \neq P$ in N, there exists one and only one geodesic (in N) joining P and Q. Let $u = u(r, \phi)$, $v = v(r, \phi)$ be the point reached at a distance r from P along a geodesic determined by the angle ϕ from some fixed direction. Then $u = u(r, \phi)$, $v = v(r, \phi)$ are of class C^{n-1} on some circle $0 \leqq r < r_0$, $0 \leqq \phi < 2\pi$ (and have a non-vanishing Jacobian if $r \neq 0$.).*

THEOREM 2*. *The element of arc-length in the (r, ϕ)-coordinates is*

$$ds^2 = dr^2 + g d\phi^2, \tag{3}$$

where $g = g(r, \phi)$ is of class C^{n-2} (and such that $g(0, \phi) \equiv 0$, while $g_r(r, \phi)$ exists and satisfies $g_r(0, \phi) \equiv 0$ even if $n = 2$).

Similar theorems hold for geodesic parallel coordinates.

2. An example of non-uniqueness. It will first be shown that the solutions of (2) need not be unique. To this end, let $h(v)$ be a positive function of class C^1 (for sufficiently small $| v |$). Choose $E(u, v) = G(u, v) = h(v)$ and $F(u, v) \equiv 0$, so that (1) becomes $ds^2 = h(du^2 + dv^2)$. If u is the independent variable, then the system of equations (2) reduces to the single equation

$$2d^2v/du^2 = (1 + (dv/du)^2)H_v, \text{ where } H = \log h, \tag{4}$$

(the subscript v denotes differentiation with respect to v); cf., e. g., the last equation on p. 403 in Darboux [1]. It will be shown that the function $h(v)$ can be chosen so that both $v \equiv 0$ and $v = u^3$ are solutions of (4).

In order that $v \equiv 0$ be a solution, it is sufficient to choose $h(v)$ so as to satisfy $H_v(0) = 0$ (i. e., $h_v(0) = 0$). On the other hand, if (4) is multiplied by dv/du, a quadrature shows that $H(v) = \log(1 + v_u^2)$ or $h(v) = 1 + v_u^2$. If $v = u^3$, then $v_u = 3u^2 = 3v^{2/3}$. In order that $v = u^3$ be a solution of (4), it is sufficient to choose $h(v) = 1 + 9v^{4/3}$. This function $h(v)$ also satisfies $h_v(0) = 0$, and so both $u = 0$ and $u = v^3$ are solutions of (4).

3. Proof of Theorem 1. Let $X = X(u, v)$ be a surface of class C^2 defined in a neighborhood of $(u, v) = (0, 0)$. It will be shown that the initial

data $(u, v) = (0, 0)$, $(u', v') = (1, 0)$ determine a unique geodesic. Suppose, if possible, that there are two distinct geodesics satisfying these initial conditions. On both of them, the variable u can be used as parameter. Let them be given by $v = v_1(u)$ and $v = v_2(u)$, respectively, for small $|u|$. Since u can be replaced, if necessary, by $-u$, it can be supposed that $v_1(u) \not\equiv v_2(u)$ for some small positive u. It can also be supposed that

$$E(0, 0) = G(0, 0) = 1, \qquad F(0, 0) = 0. \tag{5}$$

If $\mu > 0$ is sufficiently small, consider the closed curve $S = S_\mu$ which goes from $(0, 0)$ to $(\mu, v_1(\mu))$ along the first geodesic $(v = v_1(u))$, then to $(\mu, v_2(\mu))$ along the parameter curve $u = \mu$, then back to $(0, 0)$ along the second geodesic $(v = v_2(u))$. Although S need not be a simple closed curve, it is clear that one can define the interior $T = T_\mu$ of S, and that Green's theorem is applicable. Hence, the theorem of Gauss-Bonnet gives

$$\int_S \kappa\, ds + \iint_T K\, d\sigma + \pi + \alpha + \beta = 2\pi,$$

where $K = K(u, v)$ is the Gaussian curvature of the surface at (u, v); κ denotes the geodesic curvature along S; ds is the element of arc-length along S; $d\sigma = (EG - F^2)^{\frac{1}{2}} du\, dv$ is the element of area on T; finally π, α, β are the oriented exterior angles at the respective vertices $(0, 0)$, $(\mu, v_1(\mu))$, $(\mu, v_2(\mu))$ of S. Cf. the proof of the Gauss-Bonnet theorem given by van Kampen [4].

Since the geodesic curvature along geodesics is 0, the last equation reduces to

$$(\pi - \beta) - \alpha = \int_{v_1(\mu)}^{v_2(\mu)} \kappa G^{\frac{1}{2}}(\mu, v)\, dv + \int_0^{\mu} \int_{v_1(u)}^{v_2(u)} K(EG - F^2)^{\frac{1}{2}} dv\, du. \tag{6}$$

Put

$$m(\mu) = \int_0^{\mu} |v'_2(u) - v'_1(u)|\, du. \tag{7}$$

Then $m(\mu) \not\equiv 0$ for small μ and $m(\mu) \geqq |v_2(u) - v_1(u)|$ for $0 \leqq u \leqq \mu$. It is clear that the integrands in (6) are continuous (hence bounded) functions of μ, u, v near $(\mu, u, v) = (0, 0, 0)$; consequently

$$(\pi - \beta) - \alpha = O(m(\mu)). \tag{8}$$

Since $v'_1(0) = v'_2(0) = 0$, both $\alpha = \alpha(\mu)$ and $\pi - \beta = \pi - \beta(\mu)$ tend to the angle between the parameter curves at $(0, 0)$. This angle is $\pm \frac{1}{2}\pi$, since $F(0, 0) = 0$. By interchanging, if necessary, the numeration of v_1 and v_2, it can be supposed that this angle is $+\frac{1}{2}\pi$. Hence,

$$\alpha - (\pi - \beta) \sim \cos(\pi - \beta) - \cos\alpha \text{ as } \mu \to 0. \tag{9}$$

Since α is the angle between the vector $X_u(\mu, v_1(\mu)) + X_v(\mu, v_1(\mu))v'_1(\mu)$ and $X_v(\mu, v_1(\mu))$, it is seen that

$$\cos\alpha = (F + Gv'_1)\{E + 2Fv'_1 + Gv'^2_1\}^{-\frac{1}{2}}G^{-\frac{1}{2}}, \tag{10}$$

where the arguments of E, F, G are $(\mu, v_1(\mu))$ and that of v'_1 is μ. It is also seen that $\cos(\pi - \beta)$ is given by a similar expression, in which $v_1(\mu)$, $v'_1(\mu)$ are replaced by $v_2(\mu)$, $v'_2(\mu)$, respectively.

For a moment, consider μ, $v_1(\mu)$, $v'_1(\mu)$ as independent variables in the expression on the right of (10). The partial derivative of this function with respect to v'_1 at $(\mu, v_1, v'_1) = (0, 0, 0)$ is 1, by (4). Furthermore, since E, F, G are of class C^1, the expression (10) possesses a continuous (hence bounded) partial derivative with respect to $v_1 = v_1(\mu)$. Hence, as $\mu \to 0$,

$$\cos(\pi - \beta) - \cos\alpha = (1 + o(1))(v'_2(\mu) - v'_1(\mu)) + O(v_2(\mu) - v_1(\mu)). \tag{11}$$

Then (7), (8), (9) imply that $v'_2(\mu) - v'_1(\mu) = O(m(\mu))$. By (7), this means that $m'(\mu) = O(m(\mu))$. But this contains a contradiction, since $m(\mu) \not\equiv 0$ and $m(0) = 0$. The contradiction establishes Theorem 1.

4. Proof of Theorem 2. As above, let the surface $X = X(u, v)$ be of class C^2 in a vicinity of $(u, v) = (0, 0)$ and let P be the point $(0, 0)$. Assume the normalization (5) for E, F, G. According to Theorem 1, there is a unique geodesic $u = u(r, \phi)$, $v = v(r, \phi)$ satisfying the initial conditions $(u(0, \phi), v(0, \phi)) = (0, 0)$ and $(u_r(0, \phi), v_r(0, \phi)) = (\cos\theta, \sin\theta)$, where r denotes arc-length on the geodesic. The functions $u = u(r, \phi)$, $v = v(r, \phi)$ possess continuous derivatives of second order with respect to r, on some circle $0 \leqq r < r_0$ and $0 \leqq \phi < 2\pi$, since they are solutions of (2) and the solutions of (2) are uniquely determined by their initial conditions. It will first be shown that u, v and u_r, v_r are of class C^1 on some circle $0 \leqq r < r_0$, $0 \leqq \phi < 2\pi$.

Let $0 < \delta < \frac{1}{2}\pi$, then there exists a positive number $u_0 = u_0(\delta)$ such that, if $|\phi| \leqq \delta$, the equation $u = u(r, \phi)$ has a solution for $r = r(u, \phi)$ whenever $|u| \leqq u_0$. Let $v = v(u; \phi)$ denote the result of substituting $r = r(u, \phi)$ into $v = v(r, \phi)$; so that, for fixed ϕ, the arc $v = v(u; \phi)$ is the geodesic determined by the initial conditions $v(0; \phi) = 0$ and $v_u(0; \phi) = \tan\phi$. The proof of Theorem 1 will be modified to show that $v_u(u; \phi)$ (hence $v(u; \phi)$) has a continuous partial derivative with respect to ϕ for $|\phi| \leqq \delta$ and $|u| \leqq u_0$ (when $u_0 > 0$, $\delta > 0$ are sufficiently small). The arc-length, $r = r(u, \phi)$, on a geodesic is given by

$$(12) \qquad r = r(u, \phi) = \int_0^u \{E + 2Fv_u(\mu; \phi) + Gv_u^2(\mu; \phi)\}^{\frac{1}{2}} d\mu,$$

where the arguments of E, F, G are $(\mu, v(\mu; \phi))$, and accordingly is of class C^1.

Since $u = u(r, \phi)$ is determined by the equation $r = r(u, \phi)$ and since $v = v(r, \phi)$ is obtained by substituting $u = u(r, \phi)$ into $v(u; \phi)$, it will follow that $u(r, \phi)$, $v(r, \phi)$ and $u_r(r, \phi)$, $v_r(r, \phi)$ are of class C^1 on some domain $0 \leqq r < r_1$ and $|\phi| \leqq \delta$. A similar procedure can then be applied on domains $0 \leqq r < r_1(\phi_0)$ and $|\phi - \phi_0| \leqq \delta(\phi_0)$, where ϕ_0 is arbitrary. Finally, an application of the Heine-Borel theorem shows that $u_r(r, \phi)$, $v_r(r, \phi)$ are of class C^1 on some circle $0 \leqq r < r_0$ and $0 \leqq \phi < 2\pi$.

Let ϕ be fixed in such a way that $|\phi| \leqq \delta$, and let $\Delta\phi$ satisfy $|\phi + \Delta\phi| \leqq \delta$ and $|\mu| \leqq \mu_0$. In order to use the notation of the last section, put $v_1(u) = v(u; \phi)$ and $v_2(u) = v(u; \phi + \Delta\phi)$. Consider the closed curve S which goes from $(0, 0)$ to $(\mu, v_1(\mu))$ along the geodesic $v = v_1(u)$, then to $(\mu, v_2(\mu))$ along the parameter curve $u = \mu$, then back to $(0, 0)$ along the geodesic $v = v_2(u)$. Let T denote the "interior" of this curve. As above, the theorem of Gauss-Bonnet is applicable. The equation corresponding to (6) is

$$(13) \qquad (\pi - \beta) - \alpha = -\Delta\phi + \int_{v_1(\mu)}^{v_2(\mu)} \kappa G^{\frac{1}{2}} dv + \int_0^{\mu} \int_{v_1(u)}^{v_2(u)} K(EG - F^2)^{\frac{1}{2}} dv\, du,$$

since the exterior angle at the vertex $(0, 0)$ is $\pi - \Delta\phi$, instead of π as above.

The quantity κ occurring in the first integral on the right of (13) is the geodesic curvature of the parameter curve $u = \mu$ at the point (μ, v). By a standard formula, $\kappa = \det(X_v G^{-\frac{1}{2}}, (X_v G^{-\frac{1}{2}})_v G^{-\frac{1}{2}}, N)$, where N is the unit normal to the surface. Hence, $\kappa G^{\frac{1}{2}}$ is $p(\mu, v)$, where $p(u, v) = G^{-1} \det(X_v, X_{vv}, N)$ is a continuous function of (u, v). Let $q(u, v) = K(EG - F^2)^{\frac{1}{2}}$. Then (13) implies that $(\pi - \beta) - \alpha$ equals

$$(14) \qquad -\Delta\phi + \Delta v(\mu)\{p(\mu, v_1) + o(1)\} + \int_0^{\mu} \Delta v(u)\{q(u, v_1(u)) + o(1)\} du,$$

where $\Delta v = v_2(u) - v_1(u) = v(u; \phi + \Delta\phi) - v(u; \phi)$, and the $o(1)$-terms tend to 0 uniformly with respect to u (for $|u| < u_0$) as $\Delta\phi \to 0$.

On the other hand, the left side of (13) can be treated as follows:

First, (9) can be replaced by

$$\alpha - (\pi - \beta) = \{\cos(\pi - \beta) - \cos\alpha\}\{\sin\alpha + o(1)\}^{-1}.$$

The $o(1)$-term refers to $\Delta\phi \to 0$ and is uniform with respect to u for $|u| \leqq u_0$; also $\sin\alpha = \det(X_u + X_v v_{1u}, X_v, N)/|X_u + X_v v_{1u}||X_v|$. Thus $\sin\alpha = (EG - F^2)^{\frac{1}{2}}\{E + 2Fv_{1u} + Gv_{1u}^2\}^{-\frac{1}{2}}G^{-\frac{1}{2}}$, where the arguments of E, F, G are $(\mu, v_1(\mu))$.

Next, (11) can be replaced by

$$\cos(\pi - \beta) - \cos\alpha = (a + o(1))\Delta v_u + (b + o(1))\Delta v$$

where $a = a(\mu, v_1(\mu), v_{1u}(\mu))$, $b = b(\mu, v_1(\mu), v_{1u}(\mu))$ are the partial derivatives of the expression on the right of (10) with respect to v'_1 and v_1, respectively, at $(\mu, v_1, v'_1) = (\mu, v_1(\mu), v_{1u}(\mu))$. Consequently, $(\pi - \beta) - \alpha$ equals

$$(15) \qquad -\{\sin\alpha(\mu) + o(1)\}^{-1}\{(a + o(1))\Delta v_u + (b + o(1))\Delta v\}.$$

It is clear that neither $\sin\alpha$ nor the function a vanishes if the numbers u_0, δ are sufficiently small.

On interchanging the roles of μ and u, the equality of (14) and (15) can be expressed as

$$\Delta v_u/\Delta\phi = A + B\Delta v/\Delta\phi + C\int_0^u D(\mu)\Delta v/\Delta\phi\, d\mu + o(1),$$

an integro-differential equation for $\Delta v/\Delta\phi$, where A, B, C, D are continuous functions of $u, v(u;\phi)$ and $v_u(u;\phi)$ and, for fixed ϕ, are considered as functions of u. (Actually, $A = \sin\alpha/a$; $B = -(p\sin\alpha + b)/a$; $C = -A$; $D = q$.) It follows by a slight modification of the arguments occurring in the theory of differential equation (cf., e. g., [3], pp. 145-146 or 155-158), that the partial derivative $v_\phi(u, \phi)$ exists, is a continuous function and satisfies the integro-differential equation

$$(16) \qquad y' = A + By + C\int_0^u D(\mu)y(\mu)\, d\mu, \qquad (' = d/du),$$

provided that (16) has a unique solution satisfying $y(0) = 0$. On the other hand, it is easy to show (for example, by the method of successive approximations) that if A, B, C, D are continuous functions of u for $|u| \leqq u_0$, then the linear integro-differential equation (16) possesses, for $|u| \leqq u_0$, a unique solution $y = y(u)$ satisfying an arbitrary initial condition.

Hence $v_\phi(u;\phi)$ exists, is continuous and satisfies (16). But (16) implies that $v_{\phi u}(u, \phi)$ exists and is continuous. Consequently, $v_u(u;\phi)$ is of class C^1 for $|u| \leqq u_0$, $|\phi| \leqq \delta$ (if u_0, δ are sufficiently small). This completes the proof of the statement that $u = u(r, \phi)$, $v = v(r, \phi)$ and their partial

derivatives of first order with respect to r are of class C^1 on some circle $0 \leqq r < r_0$, $0 \leqq \phi < 2\pi$.

The proof also shows that if the surface $X(u, v)$ is of class C^n, where $n \geqq 2$, then $u(r, \phi)$, $v(r, \phi)$ (and their partial derivatives with respect to r) are of class C^{n-1}.

Theorem 1 and the standard arguments (cf. [1], p. 408) show that $u = u(r, \phi)$, $v = v(r, \phi)$ are functions of the normal coordinates $\xi = r\cos\phi$, $\eta = r\sin\phi$, say $u = U(\xi, \eta)$, $v = V(\xi, \eta)$. It is easily verified that U, V are of class C^1 on the circle $\xi^2 + \eta^2 \leqq r_0^2$. For example, if $\xi^2 + \eta^2 \neq 0$, then $U_\xi(\xi, \eta) = (\xi/r)u_r - (\eta/r^2)u_\phi$; while $u_r = \cos\phi + O(r)$ and $u_\phi = -r\sin\phi + o(r)$, uniformly in ϕ, as $r \to 0$. Hence, U_ξ exists, is continuous at $(\xi, \eta) = (0, 0)$ and has the value 1 there. Similarly, $U_\eta(0, 0) = 0$; $V_\xi(0, 0) = 0$; $V_\eta(0, 0) = 1$. Consequently, the Jacobian $\partial(U, V)/\partial(\xi, \eta)$ is 1 at $(0, 0)$. This proves Theorem 2.

5. Proof of Theorem 2*. Let $X = Y(r, \phi)$ denote the result of substituting $u = u(r, \phi)$, $v = v(r, \phi)$ into $X = X(u, v)$. Clearly, the scalar product $Y_r \cdot Y_r$ is identically 1, since r is arc-length along the geodesic, $\phi = \text{const.}$

It will be shown that $Y_r \cdot Y_\phi \equiv 0$. Under standard restrictions, this is merely the theorem of Gauss concerning orthogonal trajectories to the geodesics issuing from a point. In order to prove it under the assumptions at hand, the standard arguments (cf. [1], p. 410) will be suitably modified.

First, Y_{rr}, $Y_{r\phi}$ exist and are continuous, and so $(Y_r \cdot Y_\phi)_r = Y_{rr} \cdot Y_\phi + Y_r \cdot Y_{r\phi}$ exists. But $Y_{rr} \cdot Y_\phi \equiv 0$, since Y_{rr}, the principal normal to the geodesic $\phi = \text{const.}$, is parallel to the surface normal N and therefore orthogonal to Y_ϕ. Also, $Y_r \cdot Y_r \equiv 1$ implies $Y_r \cdot Y_{r\phi} \equiv 0$. Hence, $(Y_r \cdot Y_\phi)_r \equiv 0$, so that $Y_r \cdot Y_\phi$ is independent of r. On the other hand, the identity $Y_\phi(0, \phi) \equiv 0$ implies $Y_r \cdot Y_\phi \equiv 0$.

Hence, in terms of dr and $d\phi$, the element of arc-length has the form (3), where $g = Y_\phi \cdot Y_\phi$. Clearly, $g(0, \phi) \equiv 0$. Also, $g_r(r, \phi)$ exists and equals $2Y_\phi \cdot Y_{r\phi}$; so that $g_r(0, \phi) \equiv 0$. Since $g(r, \phi)$ has the smoothness asserted, the Theorem 2* is proved.

6. Remark. Using the notation of Section **5**, let the surface $X = X(u, v)$ be of class C^2 in a vicinity of $(u, v) = (0, 0)$, and let $X = Y(r, \phi)$ denote the parametrization of the surface in geodesic polar coordinates (at $(u, v) = (0, 0)$). As mentioned above (Section **1**), the Theorema Egregium is valid on surfaces of class C^2. It follows that if S is any (smooth) Jordan

curve in the circle $0 \leqq r < r_0$ such that its interior T does not contain the point $r = 0$, then

$$\tfrac{1}{2} \int_S 0\, dr + (g_r/g^{\frac{1}{2}})\, d\phi = -\int\int_T K g^{\frac{1}{2}} dr\, d\phi.$$

(This equation is merely an "integrated" form of the Frobenius formula for the Gaussian curvature when (3) holds; this formula is valid on surfaces of class C^2, as can be seen by approximating $X = X(u, v)$ suitably by smooth functions; cf. [2], Section 7.)

Consider curves S which are the boundaries of regions of the type $(0 <) r_1 \leqq r \leqq r_2$, $\phi_1 \leqq \phi \leqq \phi_2$. If the last formula line is divided by $(r_2 - r_1)(\phi_2 - \phi_1)$, and if $\phi_1, \phi_2 \to \phi$, finally, $r_1, r_2 \to r$ $(\neq 0)$, then it is seen that $g_{rr}(r, \phi)$ exists and is continuous if $r \neq 0$, and that $\partial^2 g^{\frac{1}{2}}/\partial r^2 + K g^{\frac{1}{2}} = 0$ if $r \neq 0$.

Although $g(r, \phi)$ is smoother, as a function of r, than is to be expected, it remains an open question as to whether or not $g(r, \phi)$ possesses a partial derivative with respect to ϕ (which is to be expected and hoped for; otherwise, in terms of r, ϕ, it is impossible to consider the usual equations for geodesics). It is easily verified that the orthogonal trajectory $r = \text{const.}$ has a parametric representation of class C^2 in terms of its arc-length; however, it is not clear that its parametric representation in terms of ϕ is just as smooth.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] G. Darboux, *Leçons sur la théorie générale des surfaces*, vol. II, Paris (1894).

[2] P. Hartman and A. Wintner, "On the classical embedding problem in differential geometry," *American Journal of Mathematics*, vol. 72 (1950), pp. 553-564.

[3] E. Kamke, *Differentialgleichungen reeller Funktionen*, Leipzig (1930).

[4] E. R. van Kampen, "The theorems of Gauss-Bonnet and Stokes," *American Journal of Mathematics*, vol. 60 (1938), pp. 129-138.

[5] H. Weyl, "Ueber die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, vol. 61 (1916), pp. 40-72.

ON THE HÖLDER RESTRICTIONS IN THE THEORY OF PARTIAL DIFFERENTIAL EQUATIONS.*

By AUREL WINTNER.

1. In the following considerations, which deal with solutions u, on an (x, y)-set D, of the partial differential equation

$$\Delta u + f(x, y)u = 0, \tag{1}$$

$(\Delta u = u_{xx} + u_{yy})$, everything will depend on a precise definition of "a solution u on D," as follows:

(I) The (x, y)-set D will be required to be open (and connected or, for that matter, simply-connected; either of the latter restrictions will be immaterial, since there will no loss in generality in assuming that D is "small"). The open set D will be referred to as a domain if it is not the empty set.

(II) The coefficient function, f, of (1) will be assumed to be (real-valued and) continuous on a given D.

(III) A function u, defined on D, will be called a solution, on D, of a differential equation if and only if all those derivatives of u which occur in differential equation exist, and satisfy the differential equation, at every point of D.

Remark. In the sequel it will be essential that the notion of "a u which satisfies the differential equation almost everywhere on D," a notion first used by Lichtenstein and further developed by Schauder, Leray, and others, is entirely ruled out by (I) and (III). In fact, the ejection of an unspecified zero set (or, for that matter, of an enumerable set) replaces the domain D by a set which, instead of being again a domain, can fail to contain any domain as a subset. Correspondingly, the question is *not* an issue just in a Hilbert space (such as the u-space $\operatorname{grad} u \subset L^2$ on D).

2. If (III) is applied to (1), all that follows for the smoothness of u is the existence of the derivatives u_{xx}, u_{yy} (but not that of u_{xy} and u_{yx}, and still less their equality) at every point of D. This implies the existence of u_x and u_y, but it fails to imply the existence of a complete differential du;

* Received January 24, 1950.

as a matter of fact, not even the continuity of $u(x, y)$ can be concluded. In fact, solutions of (1) which are discontinuous on a dense set are compatible with (III) even in the case, $f \equiv 0$, of Laplace's equation. This can be seen as follows:

Define on the (x, y)-plane a function $u(x, y)$ by placing $u(0, 0) = 0$, and $u(x, y) = -2xy/(x^2 + y^2)^2$ if $(x, y) \neq (0, 0)$. Then, since $u(x, 0)$ and $u(0, y)$ vanish identically, u_{xx} and u_{yy} exist and vanish, hence they satisfy $\Delta u = 0$, at the point $(x, y) = (0, 0)$. Except for this point, $u(x, y)$ is a regular harmonic function, since it is the imaginary part of $1/(x + iy)^2$. Consequently, $u(x, y)$ is a solution of $\Delta u = 0$ on D in the sense of (III), where D is the entire (x, y)-plane, although $u(x, y)$ is discontinuous at $(0, 0)$, since it is unbounded near $(0, 0)$. In order to "condense" this isolated "singularity," it is sufficient to put $u_n(x, y) = u(x - x_n, y - y_n)$ and $u^*(x, y) = \sum a_n u_n(x, y)$, where $(x_1, y_1), (x_2, y_2), \cdots$ is a sequence of points which is dense on the (x, y)-plane, and $a_1, a_2, \cdots$ is a sequence of constants, where it is understood that the elements of these two sequences must be chosen by an appropriate induction. What then results is a function u^* which, in the sense of (III), is a solution of Laplace's equation on D (with D the entire (x, y)-plane), but which is such as to be discontinuous at every point of a set of second category (in fact, $u^*(x, y)$ can be unbounded on every circle $(x - x_n)^2 + (y - y_n)^2 < \epsilon$, where $n = 1, 2, \cdots$).

In order to avoid pathological solutions of this type, only those solutions of (1) will be considered which, besides being solutions in the sense of (I)-(III), are continuous functions. This additional restriction of the smoothness required by (III) is the more natural as every continuous solution u of $\Delta u = 0$ on D is of class C'' (hence, a regular harmonic function) on D; cf. [8] or [5].

3. It is known (for references, cf. [3], pp. 1294-1295) that if D is small enough and if J is a sufficiently smooth Jordan curve contained in D, then, with reference to every smooth continuous function assigned on J, the first boundary-value problem for the interior of J (with a $u(x, y)$ which is continuous within and on J) is solvable for (1), provided that the coefficient function satisfies a Hölder condition,

$$|f(x', y') - f(x'', y'')| \leqq \text{const.}\{(x' - x'')^2 + (y' - y'')^2\}^{\frac{1}{2}\theta}, \qquad (0 < \theta < 1).$$

It will be shown that the situation becomes entirely different if this proviso is omitted (in this regard, cf. [4]) and just the continuity of the coefficient

functions is required. In fact, there will be constructed a (real-valued) $f(x, y)$ having the following property:

(*) Although $f(x, y)$ is continuous on the entire (x, y)-plane, there does not exist *any* domain D (or, equivalently, *any* circle D: $(x-x_0)^2+(y-y_0)^2<r^2$, with some (x_0, y_0) and some $r>0$) on which D the differential equation (1) possesses *some* continuous solution, $u(x, y)$, distinct from the trivial solutions, $u(x, y)\equiv 0$.

Accordingly, *every* boundary value problem belonging to (1) becomes unsolvable for *every* Jordan curve J, except when the boundary condition on J is such as to be solved by $u(x, y)\equiv 0$ (within J, hence on J). In addition, it is possible to choose a continuous $f(x, y)$ as follows:

(**) The boundary-value problem

$$u+\{\lambda+f(x, y)\}u=0, \qquad u=0 \text{ on } J, \tag{2}$$

becomes meaningless for every circle J: $(x^2-x_0)^2+(y-y_0)^2=r^2$, since it is possible that (2) has, for *no* value of the parameter λ, and on *no* (x, y)-domain, *any* continuous solution distinct from $u(x, y)\equiv 0$, even though $f(x, y)$ is continuous throughout.

This refinement, (**), of (*) will follow from the proof (*).

4. Quite another consideration, one based simply on the equation,

$$y'=-f(x, y), \qquad (' = d/dx), \tag{a}$$

of the characteristic curves, $y=y(x)$, of

$$u_x+f(x, y)u_y=0, \tag{b}$$

shows that similar possibilities can occur for linear, homogeneous, partial differential equations of *first* order also. For the inhomogeneous extension, $u_x+f(x, y)u_y=g(x, y)$, of the homogeneous equation, (b), the construction of a counter-example is more straightforward, since, according to Perron [6], p. 550, it can be made to depend on the choice of the continuous function $g(x, y)$, whilst $f(x, y)\equiv 1$, and this would not do in (b).

Only solutions of class C' will be considered for (b). Then what corresponds to the assertions of Section **3** is as follows:

(†) There exist on the (x, y)-plane real-valued, continuous functions $f(x, y)$ having the following property: There does not exist *any* domain D on which D the differential equation (4) possesses *some* solution of class C', say $u=u(x, y)$, distinct from the trivial solutions, $u(x, y)=\text{const.}$

First, if $f(x, y)$ is unspecified, and if $y=y(x)$ is any solution curve of

(a), it is clear that $du(x, y(x))/dx$ vanishes, identically in x, whenever $u(x, y)$ has a complete differential (e. g., bounded partial derivatives u_x, u_y) and is a solution of (b) (on an (x, y)-domain, D, which contains the curve $y = y(x)$). Hence, if a point (x_0, y_0) of D issues more than one characteristic, say K_1: $y = y_1(x)$ and K_2: $y = y_2(x)$, where, without loss of generality, $y_1(x) < y_2(x)$ holds for small positive $x - x_0$, then $u(x, y)$ has the same value, $u(x_0, y_0)$, at every point of $K_1 + K_2$. But it is well-known (Osgood) that if (x^*, y^*) is any point, close enough to (x_0, y_0), which is contained in the "horn" bordered by K_1 and K_2, then (a) possesses at least one solution path, K: $y = y(x)$, which joins (x^*, y^*) with (x_0, y_0). Consequently, $u(x, y)$ is constant on the entire domain of the "horn."

Choose the real-valued, continuous $f(x, y)$ in such a way that every point of the (x, y)-plane becomes a point (x_0, y_0), that is, a point of non-uniqueness, of (a) (such functions f exist; cf. [1]). Then every point (x, y) will be a boundary point of some domain $D = D_{xy}$ (viz., the open set supplied by a "horn") on which domain the function $u(x, y)$ is constant; so that $u_x \equiv 0 \equiv u_y$ on $D = D_{xy}$. Hence, if $u(x, y)$ of class C', then $u_x \equiv 0 \equiv u_y$ holds throughout. This proves (†).

4 bis. Clearly, the coefficient function, $f(x, y)$, of the example, (b), proving (†) cannot vanish identically on any (x, y)-domain. Hence, $f(x, y)$ will not vanish at some point and therefore, by continuity, on some (x, y)-square. It can therefore be assumed that $f(x, y)$ is positive (or, equivalently, that $f(x, y)$ is negative) throughout the square $(|x| \leqq 1, |y| \leqq 1)$. This leads to the following corollary of (†):

(† bis) There exist on the square $(|x| \leqq 1, |y| \leqq 1)$ continuous functions $f(x, y) > 0$, $g(x, y) < 0$ having the property that the binary linear homogeneous system

$$u_x = f(x, y)u_y + 0 \cdot v_y, \qquad v_x = 0 \cdot u_y + g(x, y)v_y$$

(which, in view of $f(x, y) \neq g(x, y)$, is a *hyperbolic* system) has on no open set contained in $(|x| \leqq 1, |y| \leqq 1)$ any solution of class C' distinct from the trivial solution: $u(x, y) = \text{const.}$, $v(x, y) = \text{Const.}$

In fact, the system in the last formula line splits into an equation for u alone and one for v alone in such a way that (†) is applicable to each of them.

5. The proof of (*) will depend on certain facts concerning the solutions of Poisson's equation,

$$\Delta u = g(x, y), \tag{3}$$

and its formal solutions,

$$(4) \qquad u_D(x,y) = (2\pi)^{-1} \iint_D g(\xi,\eta) \log r \, d\xi d\eta,$$

where $r^2 = (\xi - x)^2 + (\eta - y)^2$. The first of these facts is as follows:

(i) If $g(x,y)$ is continuous and bounded on a bounded domain D, then (3) cannot have any continuous solution on D unless the function (4) represents a solution of (3) on D.

In order to ascertain (i), let $\Delta^* v = \Delta^* v(x,y)$ denote, for any point (x,y) and for any function v defined in a vicinity of (x,y), the limit of

$$h^{-2}\{v(x+h,y) + v(x,y+h) + v(x-h,y) + v(x,y-h) - 4v(x,y)\}$$

as $h \to 0$, provided that this limit exists. It is known (cf. [8], p. 147) that if a function $v(x,y)$ is continuous on a domain D and is such that $\Delta^* v$ exists and vanishes at every point (x,y) of D, then $v(x,y)$ is a regular harmonic function on D. Furthermore, it is readily seen (cf. [8], p. 149) that if $v(x,y)$ is a function for which Δv exists at a point (x_0, y_0) (in the sense that both $u_{xx}(x_0,y_0)$ and $u_{yy}(x_0,y_0)$ exist, implying that $u_x(x,y_0)$, $u_y(x_0,y)$ exist for small $|x - x_0|$, $|y - y_0|$, respectively), then $\Delta^* v$ exists, and equals Δv, at (x_0, y_0). Finally, Zaremba [8] has shown that if $g(x,y)$ is continuous on a bounded domain D, then the function (4) satisfies, at every point of D, the following variant of (3):

$$(3\text{ bis}) \qquad \Delta^* u_D(x,y) = g(x,y)$$

(the existence of $\Delta^* u_D$ at every point of D is the essential part of the statement).

Clearly, these three facts imply the truth of (i), simply because the logarithmic potential (4) is a continuous function on D. As a matter of fact, (4) always is a function of class C' on D. Hence, the three facts leading to (i) also lead to the following lemma (cf. [8]):

(ii) If $g(x,y)$ is continuous on a domain D, and if (3) has on D a continuous solution $u(x,y)$, then $u(x,y)$ is of class C' on D.

In fact, it is clear from the proof of (i) that the situation is as follows:

(iii) If $g(x,y)$ is continuous on a bounded domain D, then (3) either does not have any continuous solution $u(x,y)$ on D or, if it has such a solution, its most general such solution is given by $u = u_D + v$, where u_D is defined by (4) and v denotes an arbitrary function which is (regular) harmonic on D.

6. The following analogue of (*) will now be proved:

(§) It is possible to choose, on the (x, y)-plane, a continuous function, $g(x, y)$, and a dense set, S, of points in such a way that, if D is any bounded domain, the function (4) fails to have a second derivative $\partial^2 u_D/\partial x^2$ at every point of the common part of D and S.

It follows from (i) that (§) actually is an analogue of (*), with (1) replaced by (3).

In the proof of (§), on which that of (*) will depend, the following criterion of Petrini [7], pp. 132-133, will be needed: If (a, b) is a point of D, and if $E(\epsilon; a, b)$ denotes the circle $r \leqq \epsilon$, where $r^2 = (\xi - a)^2 + (\eta - b)^2$, then the existence of the limit

$$\lim_{\epsilon \to 0} \iint_{D-E(\epsilon;a,b)} g(\xi, \eta) \partial^2 \log r/\partial \xi^2 \, d\xi d\eta \tag{5}$$

is necessary and sufficient in order that the function (4), where g is continuous, be such as to possess a second derivative $\partial^2 u_D/\partial x^2$ at the point (a, b) (a similar criterion holds for the existence of the other second derivatives of u_D).

Since $r^2 = (\xi - a)^2 + (\eta - b)^2$, it is clear that

$$\iint_{D-E(\epsilon;a,b)} | \partial^2 \log r/\partial \xi^2 | \, d\xi d\eta \to \infty, \text{ as } \epsilon \to 0, \tag{6}$$

holds for every (a, b). But the integral (6) is the "Lebesgue constant" which, in the "singular integral" occurring in (5), belongs to a fixed $\epsilon > 0$. It follows therefore from (6), by Lebesgue's classical construction (cf. [2], pp. 60-63), that there exists a continuous function $f(x, y) = f_{ab}(x, y)$ for which the limit (5) fails to exist. But it is well-known that Lebesgue's construction can be carried out for a sequence of points simultaneously, if the "Lebesgue constants" (belonging to the parameter, which is the above ϵ) are unbounded (as $\epsilon \to 0$) at each of the points belonging to the sequence.[1] Consequently, there exists a continuous function $g(x, y)$ having the property that the limit (5) fails to exist at each of the points $(a, b) = (a_1, b_1)$, $(a, b_2), \cdots$, where (a_n, b_n) is any point of D. Hence the assertion of (§) follows by choosing the dense set, occurring in (§), as follows:

$$S\colon (x_n, y_n), \qquad (n = 1, 2, \cdots), \tag{7}$$

[1] In fact, if the case of a single point, say P, is replaced by that of a sequence, $P_1, P_2, P_3, \cdots$, it is sufficient to replace the sequence $P, P, P, \cdots$, which occurs in Lebesgue's construction, by the sequence P_1; P_1, P_2; P_1, P_2, P_3; $P_1, \cdots$.

where (x_1, y_1), (x_2, y_2), $\cdots$ denote those points of the (x, y)-plane for which both x and y are rational.

7. Besides (§), the following criterion of comparison will be needed in the proof of (*):

(§§) Let $g^1(x, y)$, $g^2(x, y)$ be two continuous, bounded functions on a bounded domain D containing the point (a, b), and let $u_D{}^1(x, y)$, $u_D{}^2(x, y)$ denote the corresponding functions (4). Suppose that $g^1 = Gg^2$, where $G(x, y)$ is a function satisfying $G(a, b) \neq 0$ and having a complete differential $dG(x, y)$ at (a, b) (e. g., continuous derivatives G_x, G_y). Then either both or neither of the functions $u = u_D{}^1$, $u = u_D{}^2$ has a second derivative $\partial^2 u/\partial x^2$ at $(x, y) = (a, b)$.

In fact, if $r^2 = (\xi - a)^2 + (\eta - b)^2 \to 0$, the existence of dG at (a, b) means that

$$G(\xi, \eta) = G(a, b) + \text{const.}(\xi - a) + \text{Const.}(\eta - b) + o(r),$$

and implies therefore that $G(\xi, \eta) = c + O(r)$, where $c = G(a, b)$. It follows therefore from $g^1 = Gg^2$, where $g^2(\xi, \eta) = O(1)$, that

$$g^1(\xi, \eta) = cg^2(\xi, \eta) + O(r), \text{ where } c \neq 0, \tag{8}$$

since $G(a, b) \neq 0$. In addition, (8) remains true if g^1 and g^2 are interchanged and c is replaced by $1/c$. But the existence of the limit (5) is necessary and sufficient for the existence of the $\partial^2/\partial x^2$-derivative of (4) at $(x, y) = (a, b)$. It follows therefore from (8) that the assertion of (§§) will be proved if it is ascertained that

$$\iint_{D-E(\epsilon;a,b)} O(r)\partial^2 \log r/\partial\xi^2 \, d\xi d\eta \to 0, \text{ as } \epsilon \to 0.$$

But this is sure to be true if

$$\iint_{r\leqq 1} r \mid \partial^2 \log r/\partial\xi^2 \mid d\xi d\eta < \infty,$$

and the convergence of the last integral is seen by placing $\xi = a + r\cos\phi$, $\eta = b + r\sin\phi$.

8. In order to prove (*), choose a function g satisfying the requirements of (§) with regard to (3), and put $f = -g$ (or, for that matter, $f = g$). It will be shown that this f satisfies the requirements of (*) with regard to (1). This will also prove that the same f satisfies the requirements of (**) with

regard to (2) and to every $\lambda = \text{const.}$; simply because an additive constant has no bearing on the construction leading to the $g(x, y)$ of (§).

For the given $f = -g$, let u be a continuous solution of (1) on a bounded D. The assertion is that $u \equiv 0$ on D.

First, if (1) is written in the form (3), where g represents the product of the two continuous functions $-f$, u, it follows from (i)-(iii) that the given continuous solution, $u(x, y)$, of (1) must be of class C' on D, and that the logarithmic potential (4) belonging to $g = -fu$ must have a second derivative $\partial^2 u_D/\partial x^2$ at every point (x, y) of D. It follows therefore from (§§) that, if $v_D = v_D(x, y)$ denotes the logarithmic potential (4) belonging to $g = -f$, and if (x^*, y^*) is in D, then either $\partial^2 v_D/\partial x^2$ exists at (x^*, y^*) or $u(x^*, y^*) = 0$. But g was chosen in accordance with (§). Hence, $\partial^2 v_D/\partial x^2$ fails to exist at every point of an (x, y)-set, (8), which is dense on D. Consequently, u must be 0 at every point of some dense subset of D. It follows therefore from the continuity of D that $u \equiv 0$ on D.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] M. Lavrientieff, "Sur une équation différentielle du premier ordre," *Mathematische Zeitschrift*, vol. 23 (1925), pp. 97-209.

[2] H. Lebesgue, "Sur les intégrales singulières," *Annales de la Faculté des Sciences de Toulouse*, (3), vol. 1 (1910), pp. 125-127.

[3] L. Lichtenstein, "Neuere Entwicklung der Theorie partieller Differentialgleichungen zweiter Ordnung vom elliptischen Typus," *Encyklopädie der Mathematischen Wissenschaften*, II C 12 (1924), pp. 1277-1334.

[4] ———, "Neue Untersuchungen in der Theorie der linearen partiellen Differentialgleichungen zweiter Ordnung vom elliptischen Typus," *Bulletin de l'Académie Polonaise des Sciences et des Lettres*, ser. A (1931), pp. 571-598.

[5] H. Looman, "Bemerkung zur Definition der harmonischen Funktionen," *Christian Huygens*, vol. 3 (1923-24), pp. 41-42.

[6] O. Perron, "Ueber Existenz und Nichtexistenz von Integralen partieller Differentialgleichungssysteme im reellen Gebiet," *Mathematische Zeitschrift*, vol. 27 (1928), pp. 549-564.

[7] H. Petrini, "Les dérivées premières et secondes du potentiel logarithmique," *Journal de Mathématiques*, ser. 6, vol. 5 (1909), pp. 127-223.

[8] S. Zaremba, "Contribution à la théorie d'une équation fonctionnelle de la physique," *Rendiconti del Circolo Matematico di Palermo*, vol. 19 (1905), pp. 140-150.

ON THE ZETA-FUNCTIONS OF ALGEBRAIC NUMBER FIELDS II.*

By RICHARD BRAUER.

1. It was shown in the first paper[1] that if k is an algebraic number field of a fixed degree n over the field P of rational numbers, then the discriminant d of k, the regulator R of k, and the number h of classes of ideals in k are connected by the asymptotic relation

$$\log(Rh) \sim \log \sqrt{|d|} \tag{1}$$

for $|d| \to \infty$. In this note, fields k will be considered whose degree n over P is no longer fixed. It will be shown that (1) holds for every sequence of fields k normal over P for which $n/\log|d|$ tends to zero.[2] If k is an arbitrary algebraic number field and if n^* is the degree of the minimal normal field k^* over P which contains k, the set of values

$$\left(\frac{\log(Rh)}{\log\sqrt{|d|}} - 1\right)\frac{1}{n^*}$$

(for all $k \neq$ P) is bounded and has zero as its only limit point. Since $n^* \leqq n!$, this result implies (1) for fields of fixed degree n.

The procedure is similar to that of I. In particular, the method used in **4** is an extension of a method of C. L. Siegel.[3] Except for the proof of the Lemma on p. 244 of I, the present paper is independent of I.

2. We shall denote by $\rho(k)$ the residue at $s = 1$ of the zeta-function $\zeta(s, k)$ of the algebraic number field k.

LEMMA. *There exists an absolute constant c_1 such that the inequality*

$$\rho(k) \leqq c_1^n(1+\gamma)^n|d|^{1/2\gamma} \tag{2}$$

holds for all algebraic number fields k and all $\gamma \geqq 1$.[4]

---

* Received February 7, 1950; presented to the American Mathematical Society, November 25, 1949.

[1] *American Journal of Mathematics*, vol. 69 (1947), pp. 243-250. We shall quote this paper by I.

[2] As was shown by Minkowski, this quotient $n/\log|d|$ lies below a fixed constant for all algebraic number fields $k \neq$ P.

[3] C. L. Siegel, *Acta Arithmetica*, vol. 1 (1935), pp. 83-86.

[4] We shall use the notation $c_1, c_2, \cdots$ for positive constants which do not depend on the field k. Similarly $c_4(\epsilon)$ will denote a positive constant which depends on a parameter ϵ but is independent of k.

For the proof, compare Landau, *Nachrichten von der Koeniglichen Gesellschaft der Wissenschaften zu Göttingen,* Mathematisch-Physikalische Klasse 1918, pp. 478-488, proof of Lemma 1. If a is the number of real conjugates of k, and $2b$ the number of non-real conjugates, $n = a + 2b$, then the expression

$$| d |^{s/2} 2^{-bs} \pi^{-ns/2} \Gamma(s/2)^a \Gamma(s)^b \zeta(s, k) - \rho(k) \, | d |^{1/2} (2\pi)^{-2b}/s(s-1)$$

is positive for all real s. This follows from Hecke's representation of the expression as an integral.[5] The lemma is obtained by setting $s = 1 + \gamma^{-1}$ and using the inequality

$$\zeta(1 + \gamma^{-1}, k) \leqq \zeta(1 + \gamma^{-1})^n \leqq (1 + \gamma)^n$$

where $\zeta(s) = \zeta(s, \mathrm{P})$ is Riemann's zeta-function.

The value of the residue $\rho(k)$ is given by

$$\rho(k) = 2^n \pi^b Rh/w\sqrt{| d |} \tag{3}$$

where w is twice the number of roots of unity in k if $b > 0$, and $w = 2$ if $b = 0$. We need the following rough estimate of w:

$$w \leqq c_2 n^2. \tag{4}$$

For the proof of (4) we may assume $w > 2$, which implies $b > 0$. Since k contains the roots of unity of order $w/2$, we have $n \geqq \phi(w/2)$ where ϕ denotes the Euler function. Suppose now that $w > 2n^2$. Then $w/2 > n^2 \geqq \phi(w/2)^2$. Since $\phi(q)/q^{1/2} \to \infty$ for $q \to \infty$,[6] it follows that w must lie below a certain constant. This proves (4).[7]

Combination of (2), (3), and (4) shows that

$$\log(Rh) - \log\sqrt{| d |} \leqq n \log(c_1(1 + \gamma)) + (\log\sqrt{| d |})/\gamma + 2 \log n + \log c_2. \tag{5}$$

Since $n/\log | d |$ lies below a fixed constant (cf. footnote 2), it follows for $\gamma = 1$ that there exists a constant c_3 with

$$\log(Rh)/\log\sqrt{| d |} \leqq c_3 \qquad (\text{for } k \neq \mathrm{P}). \tag{6}$$

If k ranges over a sequence of fields for which $n/\log | d |$ tends to zero, then (5) implies that

$$\overline{\lim} \log(Rh)/\log\sqrt{| d |} \leqq 1 + \gamma^{-1}.$$

---

[5] See, for instance, Siegel, *loc. cit.*[3], equation (5).

[6] See, for instance, Hardy and Wright, *An Introduction to the Theory of Numbers,* Oxford 1945, Theorem 327, p. 265.

[7] If we use Landau's inequality $\phi(q) \geqq c\, q/\log\log q$ (c a positive constant, $q > 1$), (5) can be improved to $w \leqq c' n \log\log n + c''$ (c', c'' positive constants, $k \neq \mathrm{P}$).

If $\gamma \to \infty$, it follows that

$$\overline{\lim} \log (Rh)/\log\sqrt{|d|} \leqq 1 \qquad (\text{for } n/\log|d| \to 0). \tag{7}$$

Finally, if k ranges over an infinite sequence of distinct fields, then $|d| \to \infty$ and (5) yields

$$\overline{\lim}\left(\frac{\log (Rh)}{\log\sqrt{|d|}} - 1\right)\frac{1}{n} \leqq 0 \qquad (\text{for } |d| \to \infty). \tag{8}$$

3. We now prove

THEOREM 1. *Let k be an algebraic field and let K be a normal extension field of k. Denote by d the discriminant of k and by D the discriminant of K. Set $[k:\mathrm{P}] = n$ and $[K:\mathrm{P}] = N$. There exists a constant c_4 such that*

$$\rho(K)/\rho(k) \leqq c_4^{N-n}(1+\gamma)^{N-n}(|D|/|d|)^{1/2\gamma} \tag{9}$$

for all $\gamma \geqq 1$.

Proof. Let $\mathfrak{G}$ be the Galois group of K with regard to k. It was shown in I that if θ is the character of the regular representation of $\mathfrak{G}$ and if χ_0 is the 1-character of $\mathfrak{G}$, we have an equation of the form [8]

$$\theta - \chi_0 = \sum \lambda(j)\psi^*_j \tag{10}$$

where each ψ_j is an irreducible character different from the 1-character of a cyclic subgroup $\mathfrak{A}_j$ of $\mathfrak{G}$, ψ^*_j denotes the character of $\mathfrak{G}$ induced by the character ψ_j of $\mathfrak{A}_j$, and the $\lambda(j)$ are positive rational coefficients.

Let the subgroup $\mathfrak{A}_j$ of the Galois group $\mathfrak{G}$ correspond to the field Z_j; $k \subseteq Z_j \subseteq K$. Then K is cyclic over Z_j and hence it is class field to a certain group $\mathfrak{H}_j$ of ideals in Z_j. On account of the fundamental results of Artin,[9] (10) implies

$$\zeta(s,K)/\zeta(s,k) = \prod_j L(s,\psi^*_j, K/k)^{\lambda(j)}; \tag{11}$$

moreover $L(s,\psi^*_j, K/k)$ is equal to an abelian L-series:

$$L(s,\psi^*_j, K/k) = L(s,\psi_j), \tag{12}$$

where ψ_j on the right is considered as a character of the group of ideal classes (mod $\mathfrak{H}_j$) in Z_j. It follows from (11) and (12) for $s \to 1$ that

---

[8] See the Lemma on p. 244 of I. Cf. also Aramata, *Proceedings of the Imperial Academy of Japan*, vol. 9 (1933), pp. 31-34.

[9] *Abhandlungen aus dem Mathematischen Seminar Hamburg*, vol. 3 (1924), pp. 89-108; vol. 8 (1931), pp. 292-306.

$$\rho(K)/\rho(k) = \prod_j |L(1,\psi_j)|^{\lambda(j)}. \tag{13}$$

The factors on the right are finite since ψ_j is different from the 1-character.

We now need an upper estimate for the factors $|L(1,\psi_j)|$. Let $n(j)$ be the degree of Z_j over P; let $a(j)$ denote the number of real conjugates of Z_j and let $2b(j)$ denote the number of non-real conjugates. If $\mathfrak{f}_j$ is the conductor of ψ_j and if d_j is the discriminant of Z_j, we set $\kappa_j = |d_j|^{1/2}\mathfrak{N}(\mathfrak{f}_j)^{1/2}$ where $\mathfrak{N}(\mathfrak{f}_j)$ denotes the norm taken in Z_j with respect to P. Finally, we set

$$\Gamma(s,\psi_j) = \Gamma((s+1)/2)^q\Gamma(s/2)^{a(j)-q}\Gamma(s)^{b(j)},$$

where q is the number of infinite primes of Z_j appearing in $\mathfrak{f}_j$. Then, for every $s > 1$, the inequality holds [10]

$$\kappa_j 2^{-b(j)}\pi^{-n(j)/2}\Gamma(1,\psi_j)|L(1,\psi_j)| \leqq 2(\kappa_j 2^{-b(j)}\pi^{-n(j)/2})^s\Gamma(s,\psi_j)\zeta(s,Z_j).$$

Set $s = 1+\gamma^{-1}$ with $\gamma \geqq 1$. Since

$$\zeta(1+\gamma^{-1},Z_j) \leqq \zeta(1+\gamma^{-1})^{n(j)} \leqq (1+\gamma)^{n(j)},$$

we obtain an inequality of the form $|L(1,\psi_j)| \leqq c_4{}^{n(j)}(1+\gamma)^{n(j)}\kappa_j{}^{1/\gamma}$ with a positive constant c_4. Now, (13) yields

$$\rho(K)/\rho(k) \leqq \prod_j c_4{}^{n(j)\lambda(j)}(1+\gamma)^{n(j)\lambda(j)}\kappa_j{}^{\lambda(j)/\gamma}. \tag{14}$$

In order to evaluate the product on the right, we compare the number of zeros of the functions appearing in (11) in the rectangle

$$0 \leqq \Re(s) < 1, \qquad 0 < \Im(s) \leqq t \tag{15}$$

(for $t > 0$). If M_j denotes the number of zeros of $L(s,\psi_j)$ in this rectangle, then, by the analogue of the Riemann-von Mangoldt theorem,[11]

$$2\pi M_j = n(j)t\log t + (2\log\kappa_j - n(j) - n(j)\log 2\pi)t + O(\log t).$$

The same formula holds for the number of zeros of $\zeta(s,K)$ in the rectangle, if we take for $n(j)$ here the degree N of K and for κ_j the value $|D|^{1/2}$. Similarly, in the case of $\zeta(s,k)$, we must take $n(j) = n$ and $\kappa_j = |d|^{1/2}$.

---

[10] This is shown by the method used by Landau, *Mathematische Zeitschrift*, vol. 4 (1919), pp. 152-162, in the proof of Theorem 4. If we set

$$(\kappa_j 2^{-b(j)}\pi^{-n(j)/2})^s\,\Gamma(s,\psi_j)L(s,\psi_j) = \Phi(s),$$

then according to Hecke, $\Phi(s)$ can be written as a sum of two integrals $\Psi_1(s)$ and $\Psi_2(s)$. It follows from the form of these integrals that $|\Psi_i(1)|$ is at most equal to the value obtained from $\Phi(s)$ by replacing $L(s,\psi_j)$ by $\zeta(s,Z_j)$ and taking $s \geqq 1$.

[11] See Landau, *Mathematische Zeitschrift*, vol. 2 (1918), pp. 52-154, Theorem LXXI.

Move the factor $\zeta(s, k)$ in (11) to the right and raise the equation to the l-th power where l is the least common denominator of all $\lambda(j)$. Then, on comparing the number of zeros of both sides in the rectangle (15), we find

$$lNt \log t + l(\log | D | - N - N \log 2\pi)t + O(\log t)$$
$$= lnt \log t + l(\log | d | - n - n \log 2\pi)t$$
$$+ \sum_j \{t \log t\, n(j) l\lambda(j) + (2 \log \kappa_j - n(j) - n(j) \log 2\pi) l\lambda(j) t\}.$$

This yields [12]

$$N = n + \sum_j n(j)\lambda(j); \tag{16}$$

$$\log | D | = \log | d | + 2 \sum_j \lambda(j) \log \kappa_j. \tag{17}$$

Hence

$$\prod_j c_4^{n(j)\lambda(j)} (1 + \gamma)^{n(j)\lambda(j)} = c_4^{N-n} (1 + \gamma)^{N-n};$$

$$\prod_j \kappa_j^{\lambda(j)/\gamma} = (| D |/| d |)^{1/2\gamma}.$$

If these values are substituted in (14), the inequality (9) is obtained and Theorem 1 is proved.

Remark 1. In the case that K is normal over k, the lemma in **2** may be considered as a special case of Theorem 1.

Remark 2. If $K \neq k$ and we choose $\gamma = \log(| D |/| d |)$ in Theorem 1, we obtain

$$\rho(K)/\rho(k) \leqq c_5^{N-n} \log^{N-n}(| D |/| d |),$$

with a constant c_5.

4. We now have to find a lower estimate for $\rho(k)$.

THEOREM 2. *Let $\epsilon > 0$ be a positive real number. There exists a constant $c_6(\epsilon)$ depending only on ϵ such that for all fields k normal over the field* P *of rational numbers the inequality holds*

$$\rho(k) \geqq c_6(\epsilon)^{-n} | d |^{-\epsilon}. \tag{18}$$

($n = [k : \mathrm{P}]$, d the discriminant of k).

Proof. We may assume $\epsilon \leqq 1/2$.

---

[12] Instead of using the number of zeros of $L(s, \psi_j)$ in the rectangle (15), we may also compare the functional equations of both sides of (11). It is also possible to obtain (16) and (17) without analytical tools. In the case of (16), one has to use the explicit value of the $\lambda(j)$ as given in I. The formula (17) can be deduced from Artin's conductor formula, *Journal für die reine und angewandte Mathematik*, vol. 164 (1931), pp. 1-11.

Case I.[13] *For all fields k normal over* P *the function $\zeta(s, k)$ does not vanish for real s with $1 - \epsilon/n < s < 1$.*

Then $\zeta(1 - \epsilon/n, k) \leqq 0$ and a result of C. L. Siegel [14] (Lemma 1 of I) shows that

$$\rho(k) \geqq (1 - \epsilon/n)(\epsilon/n) 2^{-n} e^{-2\pi n} \mid d \mid^{-\epsilon/2n}.$$

It is evident that an inequality (18) holds in this case.

Case II. *There exists a normal field k_0 of degree n_0 over* P *such that $\zeta(s, k_0) = 0$ for a real s with $1 - \epsilon/n_0 < s < 1$.*

We may consider n_0, s, and the discriminant d_0 of k_0 as depending only on ϵ. Let K be the compositum kk_0 of the normal fields k and k_0. Then K is normal over k_0, and (11) for k_0 instead of k implies that $\zeta(s, K) = 0$ for our fixed s, since the right hand side of (11) remains finite for this s. Again, Siegel's result [14] can be applied and yields

$$\rho(K) \geqq s(1-s) 2^{-N} e^{-2\pi N} \mid D \mid^{-(1-s)/2}, \tag{19}$$

where $N = [K : \mathrm{P}]$ and where D is the discriminant of K;

$$N \leqq n_0 n; \qquad \mid D \mid \leqq \mid d \mid^{n_0} \mid d_0 \mid^{n}. \tag{20}$$

Hence

$$\mid D \mid^{-(1-s)/2} \geqq \mid D \mid^{-\epsilon/2n_0} \geqq \mid d \mid^{-\epsilon/2} \mid d_0 \mid^{-n\epsilon/2n_0},$$

and since s, n_0, d_0 depend on ϵ but not on k, (19) can be written in the form

$$\rho(K) \geqq c_7(\epsilon)^{-n} \mid d \mid^{-\epsilon/2}, \tag{21}$$

where $c_7(\epsilon)$ depends only on ϵ.

On the other hand, K is normal over k and Theorem 1 can be applied. If we choose $\gamma = n_0/\epsilon$ and take (20) into account, we obtain an inequality

$$\rho(k) \geqq \rho(K) c_8(\epsilon)^{-n} \mid d \mid^{-\epsilon/2}, \tag{22}$$

where $c_8(\epsilon)$ also depends only on ϵ. Combination of (21) and (22) yields (18) with $c_6(\epsilon) = c_7(\epsilon) c_8(\epsilon)$, and Theorem 2 is established.

It follows from (18) and (3) that an inequality of the type

$$\log (Rh) - \log \sqrt{\ } \mid d \mid \geqq - n c_9(\epsilon) - 2\epsilon \log \sqrt{\ } \mid d \mid \tag{23}$$

holds.

---

[13] In particular, we have Case I, if Riemann's conjecture is true for $\zeta(s, k)$. If $\zeta(s, k)$ does not vanish on the real line between $1 - \epsilon/\log \mid d \mid$ and 1, the same method yields the better estimate $\rho(k) \geqq c(\epsilon)^{-n} (\log \mid d \mid)^{-1}$ with a positive $c(\epsilon)$ depending only on ϵ. This result is due to Hecke.

[14] *Loc. cit.*[3], Lemma 2 on p. 85.

Since $n/\log|d|$ is bounded for all algebraic number fields $k \neq \mathrm{P}$, we see that $\log(Rh)/\log\sqrt{|d|}$ is bounded below. This together with (6) yields:

THEOREM 3. *There exists an absolute constant c_{10} such that* $|\log(Rh)| \leqq c_{10}\log\sqrt{|d|}$ *for all fields k normal over* P.

Further, if k ranges over a sequence of fields normal over k for which $n/\log|d| \to 0$, then (23) implies that $\underline{\lim}\log(Rh)/\log\sqrt{|d|} \geqq 1 - 2\epsilon$. Since $\epsilon > 0$ is arbitrary, the limit inferior is at least 1. By (7), the corresponding limit superior is at most 1. Hence we have

THEOREM 4. *If k ranges over a sequence of fields normal over* P *for which $n/\log|d| \to 0$, then* $\log(Rh) \sim \log\sqrt{|d|}$.

5. We now consider arbitrary algebraic number fields k. If $k = \mathrm{P}(\alpha_1)$, and if $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the conjugates of α_1, set

$$k^* = k(\alpha_1, \alpha_2, \cdots, \alpha_n).$$

Then k^* is the minimal normal field over P which contains k. The degree n^* of k^* over P satisfies the condition $n \leqq n^* \leqq n!$. We need an estimate for the discriminant d^* of k^* in terms of the discriminant d of k. Let θ_i denote the isomorphism of $k = \mathrm{P}(\alpha_1)$ on $\mathrm{P}(\alpha_i)$ which maps α_1 on α_i. Let $\mathfrak{e}_{ij}$ denote the greatest common divisor of all $\omega^{\theta_i} - \omega^{\theta_j}$ with integral $\omega \in k$; $(i, j$ fixed; $i < j)$. If we set $L_i = \mathrm{P}(\alpha_1, \alpha_2, \cdots, \alpha_i)$, $L_0 = \mathrm{P}$, the different of $L_i = L_{i-1}(\alpha_i)$ relative to L_{i-1} divides [15]

$$\prod_{j>i} \mathfrak{e}_{ij}$$

$(i = 1, 2, \cdots, n-1)$, and hence the different $\mathfrak{d}^*$ of $k^* = L_{n-1}$ relative to P divides

$$\prod_{i=1}^{n-1} \prod_{j>i} \mathfrak{e}_{ij}.$$

However, the square of this double product is the discriminant d of k. Thus, $\mathfrak{d}^{*2} \mid d$. Taking the norm in k^* relative to P, we obtain $d^{*2} \mid d^{n^*}$ and this implies

$$|d^*| \leqq |d|^{n^*/2}. \tag{24}$$

Apply Theorem 1 to the field k and the normal extension field k^*. If $\gamma = 1/2\epsilon$, $\epsilon < 1/2$, we find the inequality

$$\rho(k) \geqq \rho(k^*)c_{11}(\epsilon)^{-n^*}|d^*|^{-\epsilon} \tag{25}$$

[15] See Hilbert, "Zahlbericht," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 4 (1897), pp. 177-546; Chapter V.

with a suitable $c_{11}(\epsilon)$. On the other hand, Theorem 2 applied to k^* yields

$$\rho(k^*) \geqq c_6(\epsilon)^{-n^*} \mid d^* \mid^{-\epsilon}. \tag{26}$$

On combining (3), (25), and (26), we obtain

$$(2\pi)^n Rh/\sqrt{\mid d \mid} \geqq \rho(k) \geqq (c_6(\epsilon)c_{11}(\epsilon))^{-n^*} \mid d^* \mid^{-2\epsilon}.$$

Hence, for a suitable $c_{12}(\epsilon)$,

$$\log(Rh) - \log\sqrt{\mid d \mid} \geqq -n^* c_{12}(\epsilon) - 2\epsilon \log \mid d^* \mid$$

and, by (24),

$$\left(\frac{\log(Rh)}{\log\sqrt{\mid d \mid}} - 1\right)\frac{1}{n^*} \geqq -\frac{c_{12}(\epsilon)}{\log\sqrt{\mid d \mid}} - 2\epsilon. \tag{27}$$

Excluding a finite number of fields, we have $c_{12}(\epsilon)/\log\sqrt{\mid d \mid} \leqq \epsilon$. Thus, the left hand side of (27) is bounded below and does not possess any negative limit points, if we let k range over all algebraic number fields different from P. This, together with (8) yields

THEOREM 5. *If k ranges over all algebraic number fields different from* P *and if n^* denotes the degree over* P *of the minimal normal field k^* over* P *containing k, the set of all values*

$$\left(\frac{\log(Rh)}{\log\sqrt{\mid d \mid}} - 1\right)\frac{1}{n^*}$$

is bounded and possesses 0 *as its only limit point.*

UNIVERSITY OF MICHIGAN.

ON COMPLETENESS IN THE SENSE OF ARCHIMEDES.*

By L. W. COHEN and CASPER GOFFMAN.

An archimedean extension of an ordered abelian group G is a group $H \supset G$ such that if $h > \theta$, the identity, in H, then there are $g > \theta$ in G and an integer m such that $mg \geqq h$ and $mh \geqq g$. The group G is called complete in the sense of Archimedes or a-complete if it has no proper archimedean extension.

Topological or t-completeness may be defined for G in the manner of Cantor or, equivalently, in the manner of Dedekind [1, 2].[1] It has been shown [2] that if G is a-complete it is t-complete but that the converse is not true. The principal result of this paper is that G is a-complete if and only if, for every proper isolated subgroup $I \subset G$, the ordered factor group G/I is t-complete and non-discrete.

A group G is called non-discrete if to $g > \theta$ in G there is a $g' \varepsilon G$ such that $g > g' > \theta$. A subgroup $I \subset G$ is called isolated [3] if $|g| \leqq g' \varepsilon I$ implies $g \varepsilon I$. The subgroup consisting of θ alone is isolated and, in order to avoid confusion $I = [\theta]$ shall be referred to as degenerate when it occurs in the discussion. It is easy to see that if I is isolated the cosets $I_g = I + g = [g' + g \mid g' \varepsilon I]$ are segments of G and that the factor group G/I is ordered by the condtion $I_g \geqq I$, if $g \geqq \theta$ in G. The topology of a group G is determined by a limiting ordinal $\xi^* = \xi^*(G)$ with the property that if $\xi_\eta < \xi^*$ for $\eta < \eta^* < \xi^*$ then $\sup[\xi_\eta \mid \eta < \eta^*] < \xi^*$ and a strictly decreasing sequence $g_\xi \downarrow \theta$, $\xi < \xi^*$, in G. The neighborhoods in G are sets $U_\xi(g) = [g' \mid |g' - g| < g_\xi]$, $\xi < \xi^*$, $g \varepsilon G$.

LEMMA 1. *If G is non-discrete and t-complete, then the neighborhoods $U_\xi(\theta)$ of the identity are equivalent to* (a) *a sequence I_ξ, $\xi < \xi^*$, of isolated subgroups of G, or* (b) *the segments $S_n = [g \mid |g| < \gamma_n]$, $n < \omega$, of an isolated subgroup $I \subset G$ which is isomorphic to the additive group of real numbers.*

Proof. Consider the sequence $g_\xi \downarrow \theta$ in G. Then either (α) for each $\xi < \xi^*$ there is a g_{η_ξ} such that for all integers m, $mg_{\eta_\xi} < g_\xi$, or (β) there is a $\xi < \xi^*$ such that for all g_η, $\xi \leqq \eta < \xi^*$, there is an integer m_η such that $m_\eta g_\eta \geqq g_\xi$.

* Received September 17, 1949.

[1] Numbers in square brackets refer to the bibliography at the end of the paper.

Suppose (α) holds. Let I_ξ be the set of g such that for some $g_1, \cdots, g_n \in U_{\eta_\xi}(\theta)$ and integers $m_1, \cdots, m_n$, $|g| \leqq \sum_{i=1}^{n} m_i g_i$. Then $|g| \leqq g' \in I_\xi$ implies $g \in I_\xi$, and $g', g'' \in I_\xi$ implies $g' \pm g'' \in I_\xi$. Hence I_ξ is an isolated subgroup of G. Further, $U_{\eta_\xi}(\theta) \subset I_\xi \subset U_\xi(\theta)$, $\xi < \xi^*$. Since $g \in U_{\eta_\xi}(\theta)$ implies $|g| \in U_\eta(\theta)$, the first inclusion holds. Now suppose $g \in I_\xi$. Then $|g| \leqq \sum_{i=1}^{n} m_i g_i \leqq M g_0$, where $M = \sum_{i=1}^{n} |m_i|$, $g_0 = \max[|g_1|, \cdots, |g_n|] \in U_{\eta_\xi}(\theta)$. From (α) it follows that $M g_0 \in U_\xi(\theta)$. Hence $g \in U_\xi(\theta)$ and the second inclusion follows. Hence (α) implies (a). Suppose (β) holds. Let $\gamma_0 = g_\xi$. Since $2g_\eta \downarrow \theta$, there is an ω-subsequence γ_n of g_η, $\xi \leqq \eta < \xi^*$, such that $2\gamma_n < \gamma_{n-1}$. Now $\omega\text{-}\lim \gamma_n = \theta$. If not, there is a g_ζ, $\xi < \zeta < \xi^*$, such that $g_\zeta < \gamma_n$, $n < \omega$. Then for each m, $m g_\zeta < m\gamma_m < 2^m \gamma_m < \gamma_0 = g_\xi$. This contradicts (β). Hence $\xi^*(G) = \omega$ and $\lim \gamma_n = \theta$. From (β) it follows that for each $n < \omega$ there is an integer $m_n \geqq 1$ such that $m_n \gamma_n < \gamma_{n-1} < (m_n + 1)\gamma_n$. Now consider all g such that $\theta \leqq g < \gamma_0$. For each n there is a unique integer $j_n = j_n(g)$ such that $0 \leqq j_n \leqq m_n$ and $\sum_{i=1}^{n} j_i \gamma_i \leqq g < \gamma_n + \sum_{i=1}^{n} j_i \gamma_i$. Since $\theta \leqq g - \sum_{i=1}^{n} j_i \gamma_i < \gamma_n$, $\lim \sum_{i=1}^{n} j_i \gamma_i = g$. Now let $x_0 = 1$, $x_n = x_{n-1}/(m_n + 1)$, $n < \omega$. Then $\sum_{i=1}^{n} j_i x_i \leqq \sum_{i=1}^{n} x_{n-1} < 2$. It is easy to show that $x = x(g) = \sum_{i=1}^{\infty} j_i x_i$ is a homeomorphism between the set $[g \mid \theta \leqq g < g_\xi] \subset G$ and the set of real numbers $[x \mid \theta \leqq x < \sum_{i=1}^{\infty} x_{i-1}]$ since G is t-complete. Let I be the isolated subgroup generated by $[g \mid \theta \leqq g < g_\xi]$. The mapping $x = x(g)$ may be extended additively to I and this extension is an isomorphism between I and the additive group of real numbers. The segments $S_n = [g \mid |g| < \gamma_n]$ are equivalent to the neighborhoods of θ in G. Hence (β) implies (b).

LEMMA 2. *If G is a non-discrete group and H is an archimedean extension of G then H is non-discrete, $\xi^*(H) = \xi^*(G)$ and, if $V_\xi(h)$ is a neighborhood system for H, $U_\xi(g) = V_\xi(g) \cap G$ is a neighborhood system for G.*

Proof. H is non-discrete. If not, there is a smallest positive $h \in H$ and $h < g_\xi$, $\xi < \xi^*(G)$. Since H is an archimedean extension of G, there is a $g > \theta$ in G such that $g \leqq mh$ for some integer m. Then $\theta < g \leqq mh < m g_\xi$, $\xi < \xi^*(G)$ contrary to $\lim m g_\xi = \theta$ in G.

Now consider $h_\eta \downarrow 0$, $\eta < \xi^*(H)$, in H. For each h_η there is a $g_{\xi_\eta} \leqq h_\eta$. Because of the property of $\xi^*(G)$ mentioned above and the decreasing character of the g_ξ, the g_{ξ_η} can be chosen so that $\eta_1 < \eta_2 < \xi^*(H)$ implies

$g_{\xi\eta_2} \leqq h_{\eta_2} < g_{\xi\eta_1} \leqq h_{\eta_1}$. Hence $\xi^*(H) - \lim g_{\xi\eta} = \theta$ in G as well as in H. By the minimal property of $\xi^*(H)$ and $\xi^*(G)$ ([2], p. 311), $\xi^*(H) = \xi^*(G)$.

The remainder of the lemma follows at once.

LEMMA 3. *If G is a-complete and I is a proper isolated subgroup of G, then G/I is a-complete.*

Proof. Assume that G/I is not a-complete. Then there is a proper archimedean extension $K \supset G/I$. Let H be the set of pairs (k, j), $k \in K$, $j \in I$. We define addition by $(k_1, j_1) + (k_2, j_2) = (k_1 + k_2, j_1 + j_2)$. Since I is the identity in K, and θ is the identity in $I \subset G$, (I, θ) is the identity for this addition. We define order in H by the relation $(k, j) > (I, \theta)$ if $k > I$ in K or $k = I$ and $j > \theta$ in I. With these definitions, H is an ordered abelian group. The subgroup $H' = [(I_g, j) \mid g \in G, j \in I] \subset H$ is isomorphic with G. For: G being a-complete, $mx = g$ has a solution in G for all $g \in G$ and integers $m \neq 0$ ([4], p. 197). Since I is an isolated subgroup of G, $mI = I$ for all integers $m \neq 0$. From a known theorem [2] ([5], p. 766) it follows that I is a summand of G. Hence $H' \sim G$.

There is a $k_0 \in K - G/I$ and so the pair $(k_0, \theta) \in H - H'$. Now consider $(k, j) > (I, \theta)$ in H. If $k = I$, then $(k, j) \in H' \sim G$. If $k > I$, then since K is an archimedean extension of G/I, there is an $I_g > I$ such that for some integer m, $mI_g > k$ and $mk > I_g$. Now $m(I_g, j) = (mI_g, mj) > (k, j)$ and $m(k, j) = (mk, mj) > (I_g, j)$. Hence H is a proper archimedean extension of $H' \sim G$, contrary to the hypothesis.

LEMMA 4. *If G is a-complete then G is non-discrete.*

Proof. If $g > 0$ in G then $g > h > 0$ where h is the solution of $2x = g$.

We introduce some notation for convenience. Consider an extension H of an ordered abelian group G and an isolated subgroup $I \subset G$. We denote by I' the set of $h \in H$ for which there are $g(h) \in I$ such that $|h| \leqq g(h)$. It is clear that I' is an isolated subgroup of H. The mapping $f(I_g) = I'_g$ is an order preserving isomorphism of G/I into H/I'.

LEMMA 5. *If G is a non-discrete and t-complete group and H is an archimedean extension of G, then for $h \in H - G$ there is an isolated non-degenerate subgroup $I \subset G$ such that $I'_h \cap G = 0$.*

Proof. Consider $h \in H - G$. Since G is t-complete and non-discrete, there is a neighborhood $V_\xi(h)$ such that $V_\xi(h) \cap G = 0$. G satisfies the hypothesis of Lemma 1. If (a) of Lemma 1 holds for G, then by Lemma 2, $U_\xi(\theta) = V_\xi(\theta) \cap G$ is a neighborhood of θ in G, and it contains an isolated

[2] The authors are indebted to the referee for calling this theorem to their attention.

subgroup $I_\xi \subset G$. Now $I'_\xi \subset V_\xi(\theta)$, and $(I'_\xi + h) \cap G \subset V_\xi(h) \cap G = 0$. Now suppose (b) of Lemma 1 holds for G. There is an isolated subgroup $I \subset G$, isomorphic to the additive group of real numbers, such that $U_\xi(\theta) \subset I$ and is mapped on some set of real nuumbers $[x \mid |x| < \epsilon]$, $\epsilon > 0$. Since the additive group of real numbers has no proper archimedean extension, $I' = I$. Now $(I' + h) \cap G = (I + h) \cap G = 0$ since $h \notin G$.

Suppose that H is an archimedean extension of G and that σ is the set of all isolated subgroups $I \subset G$ such that $(I' + h) \cap G = 0$. The union $K = \bigcup_\sigma I$ is an isolated subgroup of G such that $(K' + h) \cap G = 0$. We shall call K the maximal isolated subgroup with this property.

THEOREM. *An ordered abelian group G is a-complete if and only if, for every isolated proper subgroup $I \subset G$, G/I is non-discrete and t-complete.*

Proof. Suppose that the condition holds and that H is an archimedean extension of G. We show that $H = G$. If not, there is an $h_0 \in H - G$. From Lemma 5, and the above remark, there is a maximal isolated subgroup $K \subset G$ such that $(K' + h_0) \cap G = 0$ and K is not degenerate. Now $K \neq G$. If not, then since there is in $K = G$ an element $g > |h_0|$, $K' + h_0 = K' \supset K$. This contradicts $(K' + h_0) \cap G = 0$.

Now H/K' is an archimedean extension of $f(G/K)$ where $K'_g = f(K_g)$ is the isomorphism of G/K into H/K' defined previously. For: if $K'_h > K'$ in H/K', then $h > \theta$ in H. Since H is an archimedean extension of G, there is a $g > 0$ such that for some integer m, $mg > h$ and $mh > g$. Hence $f(K_{mg}) = K'_{mg} \geqq K'_h$ and $K'_{mh} \geqq K'_g = f(K_g)$. Further, $K' + h_0 \in H/K' - f(G/K)$. If not, there is $K'_g = K' + h_0 \in f(G/K)$, and so $(K' + h_0) \cap G \supset K_g \neq 0$, which is false.

Now by hypothesis G/K is non-discrete and t-complete so that Lemma 5 applies to $H/K' \supset f(G/K)$ and $K'_{h_0} \in H/K - f(G/K)$. Hence there is an isolated subgroup $\mathfrak{J} \subset f(G/K)$ such that $(\mathfrak{J}' + K') \cap f(G/K) = 0$.

If $K'_g \in \mathfrak{J}'$ then $g \in K$. If not, there is $g_0 \in G - K$ such that $K' + g_0 \in \mathfrak{J}'$. Let J be the set of $g \in G$ such that for some $g' \in K$ and integer m, $|g| \leqq g' + mg_0$. Then J is an isolated subgroup of G and $K \subset J - [g_0]$. Since K is the maximal isolated subgroup of G such that $(K' + h_0) \cap G = 0$, there is a $g_1 \in (J' + h_0) \cap G$. Now $g_1 - h_0 \in J'$, and so there are $g \in J$ and $g' \in K$ and an integer m, such that $|g_1 - h_0| \leqq |g| \leqq g' + mg_0$. Hence $K'_{|g_1 - h_0|} \leqq K'_{g'} + K'_{mg_0} = K'_{mg_0} = mK'_{g_0} \in \mathfrak{J}'$, since $\mathfrak{J}'$ is an isolated subgroup of H/K', and $K'_{g'} = K'$ is the identity in $\mathfrak{J}'$. Further, $K_{g_1 - h_0} \in \mathfrak{J}'$, $K'_{g_1} \in \mathfrak{J}' + K'_{h_0}$ and so $K'_{g_1} \notin f(G/K)$. Hence $K_{g_1} \notin G/K$, contrary to $g_1 \in G$.

Since $\mathfrak{J} \neq [K]$ by Lemma 5, $\mathfrak{J}' \neq [K']$. Hence there is $h > \theta$ in H such that $K' < K'_h \in \mathfrak{J}'$. Consider any $g > \theta$ in G. If $g \in K$, then for all

integers m, $mg \in K \subset K' < K'_h$. Hence $mg < h$ for all integers m. If $g \notin K$, $K_g \subset K'_g \notin \mathfrak{A}'$. Since for all integers m, $K'_{mh} = mK'_h \in \mathfrak{A}'$, it follows that if $g > \theta$, $K'_{mh} < K'_g$. Hence $mh < g$ for all integers m. Thus there does not correspond to some $h > \theta$ in H any $g > \theta$ in G such that, for some integer m, $mg \geqq h$ and $mh \geqq g$. This contradicts the assumption that H is an archimedean extension of G. Hence $H = G$, and the sufficiency is proved.

To establish the necessity of the condition, consider an a-complete group G. If I is a proper isolated subgroup of G, then G/I is an ordered abelian group. By Lemma 3, G/I is a-complete. Since an a-complete ordered abelian group is t-complete ([2], Theorem 9), G/I is t-complete. By Lemma 4, the a-complete group G/I is non-discrete.

We conclude with examples indicating the independence of the conditions. Let R be the group of rational valued sequences $r = \{r_n\}$ ordered lexicographically. For neighborhoods of the identity we use $U_n(\theta) = [r \mid r_1 = \cdots = r_n = 0]$. The $U_n(\theta)$ are isolated subgroups of R. R is non-discrete and t-complete, but the factor groups $R/U_n(\theta)$ are not t-complete, although they are non-discrete. R is not a-complete since R^*, the group of all real sequences $x = \{x_n\}$, is a proper archimedean extension. Let N be the group of complex numbers $n + iy$, where n is an integer and y is real, ordered lexicographically. The neighborhoods of the identity are $U_n(\theta) = [iy \mid \mid y \mid < 1/n]$. N is non-discrete and t-complete. The subgroup $I = [iy \mid y \text{ real}]$ is isolated. The factor group N/I, isomorphic to the group of real integers, is discrete and, therefore, trivially t-complete. N is not a-complete, since the ordered group C of all complex numbers is a proper archimedean extension of N.

QUEENS COLLEGE, FLUSHING, N. Y.

UNIVERSITY OF OKLAHOMA, NORMAN, OKLA.

BIBLIOGRAPHY.

[1] R. Baer, "Zur Topologie der Gruppen," *Jour. für d. reine u. angew. Math.*, vol. 160 (1929), pp. 208-226.

[2] L. W. Cohen and C. Goffman, "The topology of ordered abelian groups," *Transactions of the American Mathematical Society*, vol. 67 (1994), pp. 310-319.

[3] W. Krull, "Allgemeine Bewertungstheorie," *Jour. für d. reine u. angew. Math.*, vol. 167 (1932), pp. 160-196.

[4] F. Hausdorff, *Grundzüge der Mengenlehre*, Leipzig (1914).

[5] R. Baer, "The subgroup of elements of finite order of an abelian group," *Annals of Mathematics*, vol. 37 (1936), pp. 766-781.

ON COMPLETENESS AND CATEGORY IN UNIFORM SPACE.*

By L. W. COHEN and CASPER GOFFMAN.

The problem of the relation between completeness and category in uniform space arises from results obtained for certain topological groups. For ordered abelian groups G, two concepts of completeness have been defined: G is called t-complete if every fundamental sequence has a limit in G; G is called a-complete if it has no proper archimedean extension [1, 2].[1] We have shown [1] that if G is a-complete, it is t-complete and of the second ξ^*-category.[2] There are t-complete G which are not a-complete, but it is not known whether there are t-complete G which are of the first ξ^*-category. However there are, for each $\xi^* > \omega$, topological abelian groups which are t-complete and of the first ξ^*-category [2]. The relation between the two concepts of completeness for ordered abelian groups G is given in the theorem [3]: G is a-complete if and only if, for every isolated subgroup $I \subset G$, the factor group G/I is non-discrete and t-complete.

The property of Archimedes, requiring order for its formulation, is not applicable to topological groups in general. Thus a property corresponding to a-completeness which would imply that such a group is of the second ξ^*-category is lacking. It is the purpose of this paper to supply such a concept in a form applicable to the class of uniform spaces metrizable with distances in an ordered abelian group.

We consider the uniform spaces $S = [x]$ with neighborhood systems $\mathfrak{U} = [U_\xi(x), x \in S, \xi < \xi^*]$ satisfying the conditions:

(1) $\bigcap_{\xi<\xi^*} U_\xi(x) = (x)$;

(2) $\xi_1 < \xi_2 < \xi^*$ implies $U_{\xi_1}(x) \supset U_{\xi_2}(x)$;

(3) $\eta < \xi^*$ implies there is a $\xi(\eta)$ such that $\eta < \eta^* < \xi^*$, and if $U_{\xi(\eta)}(y) \cap U_{\xi(\eta)}(x) \neq 0$ then $U_{\xi(\eta)}(y) \subset U_\eta(x)$;

* Received November 7, 1949.

[1] Numbers in brackets refer to the appended bibliography.

[2] There is a unique ordinal ξ^* associated with an ordered abelian group G. A sequence in G is a single valued function g_ξ on $\xi < \xi^*$ to G. A group G is of the first or second ξ^*-category according as it is or is not the union of nowhere dense sets N_ξ, $\xi < \xi^*$.

(4′) $\eta^* < \xi^*$ implies $\bigcap_{\eta<\eta^*} U_{\xi_\eta}(x)$ is open, where ξ^* is a limiting ordinal such that if $\xi_\eta < \xi^*$ on $\eta < \eta^* < \xi^*$, then

(*) $$\sup [\xi_\eta \mid \eta < \eta^*] < \xi^*.$$

These conditions are satisfied by ordered abelian groups G with segments as neighborhoods. It has been shown [2] that if S is t-complete and satisfies

(4) If $U_{\xi_\eta}(x_\eta)$ decreases on $\eta < \eta^* < \xi^*$ then $\bigcap_{\eta<\eta^*} U_{\xi_\eta}(x_\eta)$ is a non-empty open set,

then S is of the second ξ^*-category.

We shall associate with S a family of "factor" spaces S^λ, analogous to the factor groups G/I mentioned above, and call S a-complete if, for all λ, S^λ is t-complete. It will be shown that if S is a-complete then S is of the second ξ^*-category. It has been shown [4] that a uniform space S is metrizable with distances in an ordered abelian group G if and only if S satisfies (1), (2), (3), (4′). Our result may now be stated as follows: A metrizable uniform space S which is a-complete is of the second ξ^*-category. Since the two concepts of completeness coincide if $\xi^* = \omega$, it would seem that the proper generalization of the classical concept of completeness is a-completeness.

We assume $\xi^* > \omega$ in the remainder of this note. In [4] it is shown that S has a neighborhood system $\mathcal{V} = [V_\xi(x)]$, equivalent to $\mathcal{U}$, which satisfies (1), (2), (3), (4′) and

(5) $V_\xi(x) \cap V_\xi(y) \neq 0$ implies $V_\xi(x) = V_\xi(y)$.

We define a neighborhood system $\mathcal{W} = [W_\xi(x)]$ where $W_\xi(x) = \bigcap_{\eta<\xi} V_\eta(x)$, $\xi < \xi^*$. It is clear that $\mathcal{W}$ is equivalent to $\mathcal{V}$ and satisfies (1), (2), (3), (4′), (5) and

(6) If $\lambda < \xi^*$ is a limiting ordinal, then $W_\lambda(x) = \bigcap_{\xi<\lambda} W_\xi(x)$.

LEMMA 1. *If $\xi < \eta < \xi^*$, then $W_\xi(y) = \bigcup [W_\eta(x) \mid x \in W_\xi(y)]$.*

Proof. Consider $z \in W_\eta(x)$ where $x \in W_\xi(y)$. Then $z \in W_\eta(x) \subset W_\xi(x) = W_\xi(y)$ by (2) and (5). Hence $\bigcup [W_\eta(x) \mid x \in W_\xi(y)] \subset W_\xi(y)$. The reversed inclusion follows from (1).

We consider any limiting ordinal $\lambda \leqq \xi^*$. Let Λ be the set of limiting ordinals $\mu \leqq \lambda$ such that there is a strictly increasing, ordinal valued function $\xi(\nu)$ on $\nu < \mu$ for which $\sup[\xi(\nu) \mid \nu < \mu] = \lambda$. Since $\lambda \in \Lambda$, Λ is not empty and so contains a first ordinal $\mu^*(\lambda)$.

LEMMA 2. *$\mu^*(\lambda)$ has property (*).*

Proof. Suppose $\mu^*(\lambda)$ does not have property (*). Then there is a $\nu^* < \mu^*(\lambda)$ and an ordinal valued function $\eta(\nu)$ on $\nu < \nu^*$ such that $\sup[\eta(\nu) \mid \nu < \nu^*] = \mu^*(\lambda)$. We define $\zeta(1) = \nu(1)$ and assume that, for $\sigma < \nu^*$, there is a strictly increasing, ordinal valued function $\zeta(\tau)$ on $\tau < \sigma$ such that $\eta(\tau) \leqq \zeta(\tau) < \mu^*(\lambda)$. Let $\zeta_0 = \sup[\zeta(\tau) \mid \tau < \sigma]$. If $\zeta_0 = \mu^*(\lambda)$, then σ is a limiting ordinal $\sigma^* < \nu^* < \mu^*(\lambda)$. If $\zeta_0 < \mu^*(\lambda)$, then we define $\zeta(\sigma) = \max[\eta(\sigma), \zeta_0] + 1$. In any case, there is a limiting ordinal $\sigma^* \leqq \nu^* < \mu^*(\lambda)$ and a strictly increasing function $\zeta(\sigma)$ on $\sigma < \sigma^*$ such that $\sup[\zeta(\sigma) \mid \sigma < \sigma^*] = \mu^*(\lambda)$.

Since $\mu^*(\lambda) \in \Lambda$, there is a strictly increasing function $\xi(\nu)$ on $\nu < \mu^*(\lambda)$ such that $\sup[\xi(\nu) \mid \nu < \mu^*(\lambda)] = \lambda$. Hence $\sup[\xi(\zeta(\sigma)) \mid \sigma < \sigma^*] = \lambda$. Since $\xi(\zeta(\sigma))$ is strictly increasing on $\sigma < \sigma^*$, $\sigma^* \in \Lambda$. Hence $\mu^*(\lambda) \leqq \sigma^*$, contrary to $\sigma^* \leqq \nu^* < \mu^*(\lambda)$, and the lemma is proved.

Now for each limiting ordinal $\lambda < \xi^*$ we define $X^\lambda = W_\lambda(x)$. We define S^λ as the set of all X^λ. For $\mu^*(\lambda)$ we fix an ordinal valued, strictly increasing function ξ_η on $\eta < \mu^*(\lambda)$ for which $\sup[\xi_\eta \mid \eta < \mu^*(\lambda)] = \lambda$. We define $W_\eta{}^\lambda(X^\lambda)$ as the set of $Y^\lambda = W_\lambda(y) \subset W_{\xi_\eta}(x)$ where $X^\lambda = W_\lambda(x)$. Finally we define for S^λ the neighborhood system $W^\lambda = [W_\eta{}^\lambda(X^\lambda), X^\lambda \in S^\lambda, \eta < \mu^*(\lambda)]$. It is consistent with these definitions to put $S^{\xi^*} = S$, $X^{\xi^*} = x$, $W_{\xi^*}(X^{\xi^*}) = W_\xi(x)$.

LEMMA 3. *$S^\lambda, \lambda < \xi^*$, satisfies (1), (2), (5) with ξ^* replaced by $\mu^*(\lambda)$.*

Proof. Condition (1) follows from (6). Condition (2) and (5) are clearly satisfied.

The concepts of fundamental sequence and t-completeness are defined for S^λ in terms of $\mu^*(\lambda)$ as in [2]. We call S a-complete (complete in the sense of Archimedes) if S^λ is t-complete for each $\lambda \leqq \xi^*$. Since $S^{\xi^*} = S$, it follows that if S is a-complete it is t-complete.

LEMMA 4. *If S is a-complete then S satisfies (4).*

Proof. Consider neighborhoods $W_{\alpha\beta}(x_\beta)$ which decrease on $\beta < \beta^* < \xi^*$. By property (*), $\sup[\alpha_\beta \mid \beta < \beta^*] = \xi < \xi^*$. If there is a $\gamma < \beta^*$ such that $\alpha_\gamma = \xi$, $\bigcap_{\beta < \beta^*} W_{\alpha\beta}(x_\beta) = W_{\alpha\gamma}(x_\gamma)$ by (2), and is open by (4′).

If there is no such γ, then ξ is a limiting ordinal $\lambda < \xi^*$. We now consider the space S^λ, the ordinal $\mu^*(\lambda)$ and the fixed function ξ_η such that $\sup[\xi_\eta \mid \eta < \mu^*(\lambda)] = \lambda$.

There exists a function $\alpha_{\beta\eta}$ strictly increasing on $\eta < \mu^*(\lambda)$ such that $\sup [\alpha_{\beta\eta} \mid \eta < \mu^*(\lambda)] = \lambda$. For there is a β_1 such that $\xi_1 < \alpha_{\beta_1}$. Assume that $\alpha_{\beta\eta}$ is strictly increasing on $\eta < \sigma < \mu^*(\lambda)$ and that $\xi_\eta \leqq \alpha_{\beta\eta}$. Then since $\mu^*(\lambda)$ is the smallest ordinal in Λ, $\sup [\alpha_{\beta\eta} \mid \eta < \sigma] = \alpha < \lambda$. Now $\max [\alpha, \xi_\sigma] < \lambda$, and so there is a $\beta_\sigma < \beta^*$ such that $\max [\alpha, \xi_\sigma] < \alpha_{\beta\sigma} < \lambda$. Transfinite induction yields $\alpha_{\beta\eta}$ strictly increasing on $\eta < \mu^*(\lambda)$ for which $\sup [\alpha_{\beta\eta} \mid \eta < \mu^*(\lambda)] = \lambda$.

To simplify notation let $\alpha_{\beta\eta} = \delta_\eta$ and $x_{\beta\eta} = y_\eta$. We define $X^\lambda{}_\eta = W_\lambda(y_\eta) \, \varepsilon \, S^\lambda$, $\eta < \mu^*(\lambda)$, and show that it is a fundamental sequence. Consider $\zeta \leqq \eta < \mu^*(\lambda)$. Then $\xi_\zeta \leqq \delta_\zeta \leqq \delta_\eta < \lambda$, and $X^\lambda{}_\eta = W_\lambda(y_\eta) \subset W_{\delta_\eta}(y_\eta) \subset W_{\delta_\zeta}(y_\zeta)$ by (2) and the decreasing character of the neighborhoods $W_{\alpha\beta}(x_\beta)$. It follows that $X^\lambda{}_\eta \, \varepsilon W^\lambda{}_\zeta \, (X_\zeta)$ for $\zeta \leqq \eta < \mu^*(\lambda)$, and so $X^\lambda{}_\eta$ is a fundamental sequence in S^λ.

Since S^λ is t-complete, $\lim X^\lambda{}_\eta = X^\lambda \, \varepsilon \, S^\lambda$. There is an $x \, \varepsilon \, S$ such that $X^\lambda = W_\lambda(x)$. We show that $W_\lambda(x) \subset \bigcap_{\beta<\beta^*} W^\lambda{}_{\alpha\beta}(x_\beta)$. If not, there is a $\beta < \beta^*$ such that $W_\lambda(x) \not\subset W_{\alpha\beta}(x_\beta)$. There is a $\zeta < \mu^*(\lambda)$ such that $\alpha_\beta < \delta_\zeta < \lambda$. There is a $\rho < \mu^*(\lambda)$ such that $\delta_\zeta < \xi_\rho < \lambda$. It follows from Lemma 1 that $W_{\delta_\zeta}(y_\zeta) = \bigcup [W_{\xi_\rho}(y) \mid y \, \varepsilon \, W_{\delta_\zeta}(y_\zeta)]$. Hence either $W_{\xi_\rho}(x) \subset W_{\delta_\zeta}(y_\zeta)$ or $x \notin W_{\delta_\zeta}(y_\zeta)$. Now, by (2), $W_\lambda(x) \subset W_{\xi_\rho}(x) \subset W_{\delta_\zeta}(y_\zeta) \subset W_{\alpha\beta}(x_\beta)$ follows from the first alternative, contrary to $W_\lambda(x) \not\subset W_{\alpha\beta}(x_\beta)$. Hence $x \notin W_{\delta_\zeta}(y_\zeta)$. From (5) it follows that $W_{\delta_\zeta}(x) \cap W_{\delta_\zeta}(y_\zeta) = 0$. Recalling that $\delta_\zeta < \xi_\rho < \lambda$, we have, for $\zeta \leqq \eta < \mu^*(\lambda)$, $W_{\xi_\rho}(x) \cap W_\lambda(y_\eta) \subset W_{\delta_\zeta}(x) \cap W_{\delta_\eta}(y_\eta) \subset W_{\delta_\zeta}(x) \cap W_{\delta_\zeta}(y_\zeta) = 0$ by (2), the increasing character of $\delta_\eta = \alpha_{\beta\eta}$ on $\eta < \mu^*(\lambda)$ and the decreasing character of $W_{\alpha\beta}(x_\beta)$ on $\beta < \beta^*$. Finally, from $W_{\xi_\rho}(x) \cap W_\lambda(y_\eta) = 0$ we have $X^\lambda{}_\eta \notin W^\lambda{}_\rho(X^\lambda), \zeta \leqq \eta < \mu^*(\lambda)$. This contradicts $\lim X^\lambda{}_\eta = X_\lambda$. Hence $\bigcap_{\beta<\beta^*} W_{\alpha\beta}(x_\beta) \neq 0$.

It remains to show that $\bigcap_{\beta<\beta^*} W_{\alpha\beta}(x_\beta)$ is open. Consider $y \, \varepsilon \bigcap_{\beta<\beta^*} W_{\alpha\beta}(x_\beta)$. From (4′) it follows that $W_{\alpha\beta}(x_\beta)$, $\beta < \beta^*$, is open and so there is a ξ_β, $\beta < \beta^*$, such that $W_{\xi_\beta}(y) \subset W_{\alpha\beta}(x_\beta)$. Again, by (4′), there is a $W_\xi(y) \subset \bigcap_{\beta<\beta^*} W_{\xi_\beta}(y) \subset \bigcap_{\beta<\beta^*} W_{\alpha\beta}(x_\beta)$, and the set is open.

Since it has been shown that $\bigcap_{\beta<\beta^*} W_{\alpha\beta}(x_\beta)$ is a non-empty open set it follows that S satisfies condition (4).

THEOREM. *If S is a-complete, S is of the second ξ^*-category.*

Proof. Since S is a-complete, it is t-complete. By Lemma 4, S satisfies (4). By Theorem 7 of [2] it follows that S, satisfying (1), (2), (3), (4) and being t-complete, is of the second ξ^*-category.

This theorem may be put in the following form.

THEOREM. *A uniform space S, metrizable with distances in an ordered abelian group, which is a-complete, is of the second ξ^*-category.*

Proof. It is shown [4] that S satisfies (1), (2), (3), (4′). The conclusion follows from the previous theorem.

QUEENS COLLEGE, FLUSHING, N. Y.
UNIVERSITY OF OKLAHOMA, NORMAN, OKLA.

BIBLIOGRAPHY.

[1] L. W. Cohen and C. Goffman, "The topology of ordered abelian groups," *Transactions of the American Mathematical Society*, vol. 67 (1949), pp. 310-319.

[2] ———, and C. Goffman, "A theory of transfinite convergence," *ibid.*, vol. 66 (1949), pp. 65-74.

[3] ———, and C. Goffman, "On completeness in the sense of Archimedes," *American Journal of Mathematics*, vol. 72 (1950), pp. 747-751.

[4] ———, and C. Goffman, "On the metrization of uniform space," *Proceedings of the American Mathematical Society*, to appear.

[5] M. Fréchet, "La notion d'uniformité et les écarts abstraits," *Comptes Rendus, Academie des Sciences*, vol. 221 (1945), p. 337.*

[6] R. Doss, "Sur la théorie de l'écart abstrait de M. Fréchet," *Bulletin des Sciences Mathématiques*, ser. 2, vol. 71 (1947), pp. 110-122.*

* We are indebted to the referee for calling our attention to [5, 6]. The common elements in the structure of the Fréchet space [6, p. 117] and the space of this paper permit us to state the following theorem: The space of Fréchet [5] is metrizable with distances in an ordered abelian group.

ON THE FUNDAMENTAL EQUATIONS OF DIFFERENTIAL GEOMETRY.*

By PHILIP HARTMAN and AUREL WINTNER.

Introduction.

If a binary, symmetric, positive-definite matrix (g_{ik}) of functions $g_{ik} = g_{ik}(u_1, u_2)$ is given on a (u_1, u_2)-domain, then the problem of embedding $ds^2 = g_{ik}\,du_i\,du_k$ into a Euclidean (x_1, x_2, x_3)-space requires the determination of three functions $x_a = x_a(u_1, u_2)$ corresponding to which $\Sigma: x_a = x_a(u_1, u_2)$ is a surface on which $dx_a dx_a$ becomes identical with ds^2. Clearly, this problem presents a system of three non-linear partial differential equations for the three unknown functions $x_a(u_1, u_2)$. But it has never been attacked directly. In fact, differential geometry considers, besides the given form $g_{ik}\,du_i\,du_k$, an auxiliary quadratic form, $h_{ik}\,du_i\,du_k$, which is required to become the second fundamental form of the unknown surface Σ. What is then required is that the three unknown functions h_{ik} should satisfy a system of three partial differential equations which contain the three given functions g_{ik}, and which consist of the Theorema Egregium and of the two equations of Mainardi-Codazzi. This system of $1 + 2$ equations, usually considered under the analyticity assumption of Cauchy-Kowalewski, was treated in [4], along with the remaining issue of the actual embedding, with such a minimum of "real" assumptions of mere differentiability as is attainable today.

In Part 1 of the present paper the "remaining issue of embedding" will be treated under assumptions which are less generous and more natural than those which have to be assumed in the classical treatment of this problem. The latter consists of solving, with given g_{ik} and given h_{ik}, the so-called derivation equations of surface theory. These are supplied by the derivation formulae of Gauss and of Weingarten; derivation formulae the system of which is a "total" system, of the Jacobi type, for the Cartesian coordinates $x_a(u_1, u_2)$ of Σ, with the $1 + 2$ partial differential equations of the Theorema Egregium and of Mainardi-Codazzi as integrability conditions.

It was pointed out in [4], §2, that the $1 + 2$ partial differential equations contain a certain type of paradox in differentiability for the three unknown functions $x_a(u_1, u_2)$, if the given functions $g_{ik}(u_1, u_2)$, $h_{ik}(u_1, u_2)$ are referred to an actual Σ of class C^n, where $n < \infty$. A corresponding paradox presents

* Received January 24, 1950.

itself in the usual treatment of the derivation formulae. This will be disposed of in Part 1, by lightening the classical assumptions of differentiability. The latter are necessitated by the usual formulation of the $1+2$ equations which represent conditions of integrability for the total system.

Accordingly, what is involved is a formulation of the integrability conditions of a total system in such a way that no superfluous assumptions of differentiability are made. This will be accomplished by using integral conditions (along arbitrary closed curves of integration), instead of the local field conditions of the standard theory of total systems.

The total system in the differential geometry of surfaces is a linear one. It will, however, be seen at the end of Part 1 that a refined formulation of the existence theorem holds for non-linear total systems also. The reduction of the condition of integrability is relevant for other applications also; cf. e. g., [8].

Part 3 has about the same bearing on the derivation formulae of the differential geometry of curves (Frenet) as Part 1 does on those of surfaces (Bonnet). It is true that the differential equations are now ordinary, and therefore nothing like the existence theorem of Part 1 is involved. Nevertheless, the usual treatment contains again a "paradox in differentiability." The latter is due to the standard definition of torsion; it will be dissolved by a more geometrical approach.

Part 2 is, in the main, independent of the rest of the paper. It deals, on the C^2-level, with the various definitions of a torse and their equivalence.

Part 1. The Fundamental Equations of Surfaces.

1. Let X denote the vector (x, y, z). The vector function $X(u, v)$, defined on a (u, v)-domain, is called a surface of class C^n, where $n \geqq 1$, if each component of X is a function of class C^n and, in addition, the vector product (X_1, X_2) does not vanish (the subscripts 1, 2 denote partial differentiation with respect to u, v, respectively). Let

$$N = (X_1, X_2)/|(X_1, X_2)| \tag{1}$$

be the unit normal vector of the surface.

On a surface of class C^2, the derivation formulae of Gauss,

$$dX_i = (\Gamma^\alpha_{i\beta} X_\alpha + h_{i\beta} N)\, du^\beta \text{ for } i = 1, 2, \tag{2}$$

and those of Weingarten,

$$dN = -g^{\alpha\beta} h_{\gamma\alpha} X_\beta du^\gamma, \tag{3}$$

must hold, if (u^1, u^2) denotes (u, v), and (g_{ij}), (h_{ij}) are the matrices of the first and second fundamental forms of the surface, (g^{ij}) is the reciprocal matrix $(g_{ij})^{-1}$, finally the Γ^i_{jk} are the Christoffel symbols of the second kind.

If X is a surface of class C^3, the integrability conditions of (2) and (3), which state that $N_{12} = N_{12}$ and that $X_{i12} = X_{i21}$ for $i = 1, 2$, are equivalent to the equations of Mainardi-Codazzi,

$$\partial h_{i1}/\partial v - \partial h_{i2}/\partial u = \Gamma^a_{i2} h_{a1} - \Gamma^a_{i1} h_{a2} \text{ for } i = 1, 2, \tag{4}$$

and to the Theorema Egregium,

$$[(g_{11v} - g_{12u})/g]_v - [(g_{12v} - g_{22u})/g]_u = -(2K + \tfrac{1}{2}D/g^4)g, \tag{5}$$

where g denotes $(\det g_{ik})^{\frac{1}{2}}$, and K the representation $\det h_{ik}/\det g_{ik}$ of the Gaussian curvature, finally D is a 3-rowed determinant in which the first column consists of g_{11}, g_{12}, g_{22}, and the second and third columns are formed by partial derivatives of g_{11}, g_{12}, g_{22} with respect to u and v, respectively.

Conversely, the classical existence theorem of Bonnet states that if the symmetric matrices $(g_{ik}(u, v))$ and $(h_{ik}(u, v))$ are of class C^2 and class C^1, respectively, and if (g_{ik}) is positive-definite, finally if the integrability conditions, (4) and (5), are satisfied, then there exist surfaces $X(u, v)$ of class C^3 for which (g_{ik}), and (h_{ik}) become the matrices of the first and second fundamental forms, respectively.

2. There is an unsatisfactory element in this theorem. First, since every surface of class C^2 (not necessarily C^3) possesses first and second fundamental forms, whose matrices (g_{ik}), (h_{ik}) are of class C^1 and C^0, respectively, it is meaningful to ask whether two given symmetric matrices, (g_{ik}) and (h_{ik}), of class C^1 and C^0, respectively, are the matrices of the first and second fundamental forms of a surface X of class C^2. But Bonnet's theorem furnishes an answer to this question only if further smoothness assumptions on the g_{ik} and h_{ik} are made; and correspondingly, in the affirmative cases, it assures the existence of surfaces X which are smoother than required by this question.

It turns out that these objections can be overcome by restating the equations (4) and (5) in an "integrated" form. The equations (4) and (5) are thus replaced by the requirement that the relations

$$\int_J h_{i1}du + h_{i2}dv = \iint_S (\Gamma^a_{i1} h_{a2} - \Gamma^a_{i2} h_{a1})\,du\,dv, \text{ where } i = 1, 2, \tag{6}$$

and

$$\int_J g^{-1}\{(g_{11v} - g_{12u})du + (g_{12v} - g_{22u})dv\} = \iint_S (2K + \tfrac{1}{2}Dg^{-4})g\,du\,dv \tag{7}$$

should hold whenever J is a Jordan curve of class C^1, and S the interior of J.

Whereas (4) and (5) need not be meaningful on a surface of class C^2, the relations (6) and (7) are meaningful and hold on every surface of class C^2. This follows from the considerations below. Incidentally, (7) shows that the Theorema Egregium, which is usually stated for surfaces of class C^3, is true for surfaces of class C^2. Actually, this statement can be deduced from the fact that the theorem of Gauss-Bonnet holds on surfaces of class C^2 (Weyl [9], pp. 42-44; cf. van Kampen, [5], p. 136).

The modified version of the theorem of Bonnet, which meets the objections mentioned above, is as follows:

(I) *Let $g_{ik}(u, v)$ and $h_{ik}(u, v)$, where $i, k = 1, 2$, be defined on a simply-connected (u, v)-domain, R. Necessary and sufficient in order that (g_{ik}), (h_{ik}) be the matrices of the first and second fundamental form of a surface $X(u, v)$ of class C^2 on R is that (g_{ik}) be symmetric, positive-definite, of class C^1, that (h_{ik}) be symmetric and continuous, and that (g_{ik}), (h_{ik}) satisfy (6) and (7) for arbitrary simple closed curves J of class C^1 in R.*

When these conditions are satisfied, $X(u, v)$ is uniquely determined up to translations and rotations of the (x, y, z)-space.

This theorem will be deduced from a modification of the standard existence theorem on systems of total differential equations. In fact, this modified theorem on systems of total differential equations leads to (I) in the same way as the standard theorem leads to Bonnet's theorem.

3. Objections, similar to those mentioned above, can be raised against the standard existence theorem on systems of total differential equations. In this theorem, it is assumed that the given data are of class C^1 and, under this hypothesis, the existence of solutions of class C^2 is proved; cf. [8], §1. It would seem more natural to consider only continuous coefficient functions, and to ask whether or not solutions of class C^1 exist. The answer to this question is given by (II) below when the system is linear. In the non-linear case, treated by (III) below, it is difficult to state integrability conditions without making some differentiability assumptions.

In what follows, the tensor sum convention, used above, will continue to be applied to Greek letters, but not to Latin letters. For example, $a_{i\alpha}z^\alpha$ is the sum $\sum_\alpha a_{i\alpha}z^\alpha$, while $a_{ij}z^j$ represents one term (not a sum).

The theorem on total differential equations needed in the proof of (I) is as follows:

(II) *Let $a^i{}_{jk}(u^1, \cdots, u^m)$ and $b^i{}_k(w^1, \cdots, u^n)$, where $i, j = 1, \cdots, m$ and $k = 1, \cdots, n$, be continuous functions on the n-dimensional cube R:*

$|u^k| \leqq r$, *where* $k = 1, \cdots, n$. *Necessary and sufficient in order that, for every set of constants* $c^1, \cdots, c^m$, *the system of total differential equations*

$$dz^i = (a^i{}_{\alpha\gamma} z^\alpha + b^i{}_\gamma)\, du^\gamma \qquad (i = 1, \cdots, m) \tag{8}$$

possess a solution $z^1(u^1, \cdots, u^n), \cdots, z^m(u^1, \cdots, u^n)$ *of class* C^1 *on* R *satisfying the initial conditions*

$$z^i(0, \cdots, 0) = c^i \text{ for } i = 1, \cdots, m, \tag{9}$$

is that the integrability conditions

$$\int_J a^i{}_{jk} du^k + a^i{}_{jl} du^l = \iint_S (a^i{}_{\alpha k} a^\alpha{}_{jl} - a^i{}_{\alpha l} a^\alpha{}_{jk})\, du^k du^l, \tag{10}$$

and

$$\int_J b^i{}_k du^k + b^i{}_l du^l = \iint_S (a^i{}_{\alpha k} b^\alpha{}_l - a^i{}_{\alpha l} b^\alpha{}_k)\, du^k du^l \tag{11}$$

hold whenever J *is a simple closed curve of class* C^1 *in* R *bounding a domain* S *in a plane parallel to the* (u^k, u^l)*-plane. When the integrability conditions* (10) *and* (11) *are satisfied, the solution of* (8) *and* (9) *is unique.*

In the proof of (II) and in the deduction of (I) from (II), the following lemma will be useful.

LEMMA. *Let* $a_1(u, v)$, $a_2(u, v)$, $c(u, v)$ *be continuous functions defined on a simply-connected* (u, v)*-domain* R *and let*

$$\int_J a_\alpha du^\alpha = \iint_S c\, du\, dv, \qquad (u^1, u^2) = (u, v), \tag{12}$$

whenever J *is a simple closed curve of class* C^1 *and* S *is its interior. Then, if* $z(u, v)$ *is a function of class* C^1 *on* R,

$$\int_J z a_\alpha du^\alpha = \iint_S (a_2 z_u - a_1 z_v + cz)\, du\, dv. \tag{13}$$

4. The derivation formulae (2) and (3) form a system of the type (8), where $m = 9$, $n = 2$ (and $b^i{}_k \equiv 0$). The proof that (6) and (7) are necessary and sufficient for the corresponding set of integrability conditions (10) can be verified as in the usual proof of the fact that, under standard conditions, (4) and (5) are equivalent to the complete set of integrability conditions (cf., e. g., [1], pp. 175-178). In the case at hand, the legitimacy of the corresponding operations (of considering linear combinations of integrability conditions) follows from the Lemma.

The verification in question can be sketched as follows: The system (2), (3) splits into three identical systems, corresponding to the three components of the vectors X_1, X_2, N. Hence, it is sufficient to consider the systems (2), (3) in which X_1, X_2, N are treated as scalars, z^1, z^2, z^3. The integrability conditions (10) to be considered are therefore satisfied. Two, corresponding to $i = 1, 2$ and $j = 3$, are identical with (6). Two, corresponding to $i = 3$ and $j = 1, 2$, involve

$$\int_J g^{ja} h_{a\beta} du^\beta, \qquad j = 1, 2,$$

and are seen to be equivalent to (6), by virtue of the Lemma, since $\det g_{ik} \neq 0$. The remaining four are of the form

$$\int_J \Gamma^j_{ik} du^a = \iint_S (\Gamma^a_{i1}\Gamma^j_{a2} - \Gamma^a_{i2}\Gamma^j_{a1} + g^{ja}h_{a1}h_{i2} - g^{ja}h_{a2}h_{i1})\, du\, dv,$$

where $i, j = 1, 2$. By virtue of the Lemma, these equations are equivalent to a set involving

$$\int_J g^{-1} g_{\alpha j} \Gamma^\alpha_{i\beta} du^\beta, \qquad i, j = 1, 2.$$

It can be verified that two in the last set of equations are trivially satisfied, and that the other two are identical with (7). This proves that the integrability conditions (10), (11) for the system (2), (3) are equivalent to (6) and (7).

The proof of (I) will be omitted, since it follows the usual proof of Bonnet's theorem, except that (II) replaces the standard theorem on systems of total differential equations.

5. The proof of the Lemma will now be given. Let (u_0, v_0) be an (interior) point of R; let $\delta > 0$ be so small that the square

$$S = S_\delta\colon\ |u - u_0| \leqq \delta, |v - v_0| \leqq \delta$$

is in R; finally, let $J = J_\delta$ be the boundary of S. In order to prove the Lemma, it is sufficient to show that if the line integral on the left of (13) is divided by δ^2, then, as $\delta \to 0$, the quotient tends to the integrand, at the point (u_0, v_0), of the double integral on the right of (13), and that this limit process is uniform with respect to (u_0, v_0) on any closed set interior to R.

Since $a_1(u, v)$, $a_2(u, v)$ are continuous and $z(u, v)$ is of class C^1, it follows that, for any point (u, v) of S,

$$a_i(u, v) = a_i^0 + o(1) \tag{14}$$

and

$$z(u, v) = z^0 + (u - u_0)z_u{}^0 + (v - v_0)z_v{}^0 + o(\delta), \tag{15}$$

where the superscript 0 on a function symbol signifies that the argument of the function is (u_0, v_0). The o-terms in (14) and (15) refer to $\delta \to 0$ and are uniform with respect to (u_0, v_0) on any closed set interior to R. Hence the line integral in (13) is the sum

$$z^0 \int_J a_\alpha du^\alpha + \int_J \{(u - u_0)z_u{}^0 + (v - v_0)z_v{}^0\}(a_\alpha{}^0 du^\alpha) + o(\delta^2).$$

By (12), the first term in this sum is, uniformly, $z^0(c^0 + o(1))\delta^2$, since $c(u, v)$ is continuous. The second term in the last formula line is $(a_2{}^0 z_u{}^0 - a_1{}^0 z_v{}^0)\delta^2$. These relations imply the assertion to be proved, and complete the proof of the Lemma.

6. Theorem (II) will now be proved. The sufficiency of (10) and (11) for the existence of a solution $z^1(u^1, \cdots, u^n), \cdots, z^m(u^1, \cdots, u^n)$ of class C^1 of the system (8) and (9) on R will be verified first. To this end, successive approximations will be employed.

On R, put

$$z^i{}_0(u^1, \cdots, u^n) = c^i \text{ for } i = 1, \cdots, m. \tag{16}$$

Suppose that the functions $z^1{}_r, \cdots, z^m{}_r$ have been defined and are of class C^1 on R. Put

$$z^i{}_{r+1}(u^1, \cdots, u^n) = \int_{(0,\ldots,0)}^{(u^1,\ldots,u^n)} (a^i{}_{\alpha\gamma} z^\alpha{}_r + b^i{}_\gamma)du^\gamma + c^i \tag{17}$$

for $i = 1, \cdots, m$, where the integral is a line integral along a path consisting of n line segments parallel to the u^1-, u^2-, $\cdots$, u^n-axis, in this order (some of the line segments can reduce to points).

It will be shown that (17) is of class C^1 in R, and that its partial derivatives are given by

$$\begin{aligned} \partial z^i{}_{r+1}/\partial u^j &= a^i{}_{\alpha j} z^\alpha{}_r + b^i{}_j \\ &+ \sum_{k=j+1}^{n} \int_0^{u_k} \{a^i{}_{\alpha k}(\partial z^\alpha{}_r/\partial u^j - a^\alpha{}_{\beta j} z^\beta{}_r - b^\alpha{}_j) \\ &- a^i{}_{\alpha j}(\partial z^\alpha{}_r/\partial u^k - a^\alpha{}_{\beta k} z^\beta{}_r - b^\alpha{}_k)\}dx, \end{aligned} \tag{18}$$

where the argument of the integrand is $(u^1, u^2, \cdots, u^{k-1}, x, 0, \cdots, 0)$. Let $P = (p^1, \cdots, p^n)$ be a point of R and let $\Delta \neq 0$ be such that

$$\Delta z^i{}_{r+1} = z^i{}_{r+1}(p^1, p^{j-1}, p^j + \Delta, p^{j+1}, \cdots, p^n) - z^i{}_{r+1}(p^1, \cdots, p^n) \tag{19}$$

is defined. If, in the expression

$$\Delta z^i{}_{r+1} + \int_{p^j+\Delta}^{p^j} (a^i{}_{aj} z^a + b^i{}_j)\, dx, \tag{20}$$

the argument of the integrand is $(p^1, \cdots, p^{j-1}, x, p^{j+1}, \cdots, p^n)$, it is clear from (17) that the expression (20) is equal to

$$\sum_{k=j+1}^{n} \int_{J_k} z^a{}_r (a^i{}_{aj} du^j + a^i{}_{ak} du^k) + \int_{J_k} b^i{}_j du^j + b^i{}_k du^k, \tag{21}$$

where J_k is the boundary of the rectangle in which $u^l = p^l$ if $l < j$ or $j < l < k$, and $u^l = 0$ if $l > k$, while u^j varies from p^j to $p^j + \Delta$ and u^k varies from 0 to p^k. Since it is assumed that $z = z^i{}_r$ is of class C^1 in R for $i = 1, \cdots, m$, the Lemma is applicable to each of the integrals in the first term of (21). Hence, the function (21) is identical with

$$\sum_{k=j+1}^{n} \int_{p^j}^{p^j+\Delta} \int_{0}^{p^k} \{(a^i{}_{ak} \partial z^a{}_r/\partial u^j - a^i{}_{aj} \partial z^a{}_r/\partial u^k) + z^a{}_r (a^i{}_{\beta j} a^\beta{}_{ak} - a^i{}_{\beta k} a^\beta{}_{aj}) + a^i{}_{aj} b^a{}_k - a^i{}_{ak} b^a{}_j\} du^j du^k, \tag{22}$$

where the argument of the integrand is $(p^1, \cdots, p^{j-1}, u^j, p^{j+1}, \cdots, p^{k-1}, u^k, 0, \cdots, 0)$. If the ratio $\Delta z^i{}_{r+1}/\Delta$ is calculated by equating expressions (20) and (22), and Δ then tends to 0, it is seen that $\partial z^i{}_{r+1}/\partial u^j$ exists at P, and that (18) holds if $(p^1, \cdots, p^n)$ is replaced by $(u^1, \cdots, u^n)$. Clearly, (18) implies the continuity of the function $\partial z^i{}_{r+1}/\partial u^j$ on R. Consequently, $z^1{}_{r+1}, \cdots, z^m{}_{r+1}$ are of class C^1 on R.

That the functions (17), where $i = 1, 2, \cdots, m$, converge uniformly, as $r \to \infty$, and that the limits, say $z^1, \cdots, z^m$, are of class C^1 on R and satisfy (8) and (9), follows by standard procedures. Furthermore, these procedures show that the solution of (8) and (9) is unique. The necessary calculations follow, without change, from those of [7], pp. 44-45 and pp. 46-49.

7. It remains to prove that the integrability conditions (10), (11) are satisfied if the existence of a solution $z^1, \cdots, z^m$ of (8) and (9), where the constants $c^1, \cdots, c^m$ are arbitrary, is assumed. This solution is necessarily unique. In fact, the successive approximations (17) are defined and continuous on R and converge uniformly on R (whether or not (10) and (11) are assumed). It follows by standard arguments, involving the use of successive approximations, that if a solution of (8) and (9) exists, it must be the limit of those successive approximations.

The difference of two solutions of (8) is a solution of the homogeneous

system obtained by placing $b^i{}_k \equiv 0$ in (8). Hence, the assumption that (8) and (9) have a solution implies that the homogeneous system and (9) have a solution. As above, this solution is necessarily unique. Consequently, if $z^1, \cdots, z^m$ is a solution of the homogeneous system and $z^1, \cdots, z^m$ vanish at some point of R, then they vanish identically. It follows that there exist m solutions of the homogeneous system, and that every solution is a linear combination (with constant coefficients) of these m solutions. Furthermore, if $c^1, \cdots, c^m$ are arbitrary constants and $(u^1{}_0, \cdots, u^n{}_0)$ is an arbitrary point of R, then there exists a solution $z^1, \cdots, z^m$ of the homogeneous system satisfying

$$z^1(u^1{}_0, \cdots, w^n{}_0) = c^1, \cdots, z^m(u^1{}_0, \cdots, u^n{}_0) = c^m. \tag{23}$$

In order to prove (10), let $(u^1{}_0, \cdots, u^n{}_0)$ be an interior point of R, and let $\delta > 0$ be so small that $J = J_\delta$, the boundary of the square

$$S_\delta: \qquad u^j = u^j{}_0 \text{ if } j \neq k, l; \ | u^k - u^k{}_0 | < \delta \text{ and } | u^l - u^l{}_0 | < \delta,$$

is in R. It is sufficient to show that if the line integral on the left of (10) is divided by δ^2, then the quotient tends, as $\delta \to 0$, uniformly to the integrand, at $(u^1{}_0, \cdots, u^n{}_0)$, of the double integral on the right of (10), and that this limit process is uniform with respect to $(u^1{}_0, \cdots, w^n{}_0)$ on any closed set interior to R.

To this end, let $z^1, \cdots, z^n$ be the solution of class C^1 of (23) and of the homogeneous system obtained from (8) by placing $b^i{}_k \equiv 0$. Then

$$\int_J z^\alpha (a^i{}_{\alpha k} du^k + a^i{}_{\alpha l} du^l) = 0, \tag{24}$$

and, on S,

$$z^\alpha = c^\alpha + (u^\gamma - u^\gamma{}_0) a^\alpha{}_{\beta\gamma 0} c^\beta + o(\delta) \quad \text{and} \quad a^i{}_{jk} = a^i{}_{jk0} + o(1),$$

where $a^i{}_{jk0} = a^i{}_{jk}(u^1{}_0, \cdots, u^n{}_0)$. The o-terms refer to $\delta \to 0$ and are uniform with respect to $(u^1{}_0, \cdots, u^n{}_0)$ on any closed set interior to R. It follows from (24) that

$$c^\alpha \int_J a^i{}_{\alpha k} du^k + a^i{}_{\alpha l} du^l = c^\beta \int_J (u^\gamma - u^\gamma{}_0) a^\alpha{}_{\beta\gamma 0} (a^i{}_{\alpha k 0} du^k + a^i{}_{\alpha l 0} du^l) + o(\delta^2).$$

But the right-hand side of this equation is

$$c^\beta (a^i{}_{\alpha l 0} a^\alpha{}_{\beta k 0} - a^i{}_{\alpha k 0} a^\alpha{}_{\beta l 0}) \delta^2 + o(\delta^2).$$

If c^α is chosen to be 1 or 0 according as $\alpha = j$ or $\alpha \neq j$, the desired limit relation follows. This proves the necessity of condition (10).

The proof of the necessity of (11) is similar and will be omitted.

8. Theorem (II) has an analogue for non-linear systems of total differential equations.

(III) *Let* $a^i{}_k(u^1, \cdots, u^n; z^1, \cdots, z^m)$ *be continuous on the parallelepiped* R: $(|z^i| \leqq b,\ i = 1, \cdots, m;\ |u^k| \leqq a,\ k = 1, \cdots, n)$ *and let* $|a^i{}_k| \leqq M$ *on* R. *Suppose further that the functions* $a^i{}_k$ *possess continuous partial derivatives with respect to the variable* $z^1, \cdots, z^m$ *on* R, *and that*

$$(25) \qquad \int_J a^i{}_k du^k + a^i{}_l du^l = \iint_S (a^a{}_l \partial a^i{}_k/\partial z^a - a^a{}_k \partial a^i{}_l/\partial z^a)\, du^k du^l$$

holds whenever J *is a simple closed curve of class* C^1 *in* R *bounding a domain* S *in a plane parallel to the* (u^k, u^l)*-plane. Then, on the parallelepiped* $|u^k| \leqq \min(a, b/nM)$, *the system of total differential equations* $dz^i = a^i{}_a du^a$ *has a unique solution* $z^1(u^1, \cdots, u^n), \cdots, z^m(u^1, \cdots, u^n)$ *of class* C^1, *satisfying the initial conditions* $z^i(0, \cdots, 0) = 0$.

(III) differs from the standard theorems in that in (III) it is not supposed that the functions $a^i{}_k$ possess partial derivatives with respect to the variables $u^1, \cdots, u^n$.

The proof of (III) will not be given, since it is similar to the proof of (II). It is only necessary to remark that the above Lemma can be replaced by the following fact: The integrability conditions (25) imply that if $z^1(u^1, \cdots, u^n), \cdots, z^m(u^1, \cdots, u^n)$ are functions of class C^1 on a (simply connected) sub-domain of R, then

$$\int_J a^i{}_k du^k + a^i{}_l du^l = \iint_S \{(a^a{}_l - \partial z^a/\partial u^l)\partial a^i{}_k/\partial z^a - (a^a{}_k - \partial z^a/\partial u^k)\partial a^i{}_l/\partial z^a\}\, du^k du^l,$$

where the argument of the integrands is $(u^1, \cdots, u^n; z^1(u^1, \cdots, u^n), \cdots, z^m(u^1, \cdots, u^n))$, and J, S have the same significance as above. (The relations (25) represent the particular case in which $z^1(u^1, \cdots, u^n), \cdots, z^m(u^1, \cdots, u^n)$ are constants.)

Part 2. On Torses.

9. Let K denote Gaussian curvature. It is well-known that, under suitable smoothness restrictions, the solutions of the fundamental equations of a surface, when $K \equiv 0$, can be characterized by either of two properties: (i) the surface is isometric with a plane; (ii) the surface is a torse. But the usual presentations either assume more than what is needed to prove the

equivalence of the three conditions (i), (ii), $K \equiv 0$ or fail to state the required restrictions. That some smoothness conditions are needed before these three properties become equivalent was shown, by a discussion of part of the problem, by Lebesgue [6], pp. 319-342.

Although the notions of a Gaussian curvature, a torse, and isometric equivalence, can be defined for surfaces of class C^1, this section will be restricted to the consideration of surfaces of class C^2. Before proceeding further, the pertinent definitions for surfaces $X(u, v)$ of class C^2 will be stated:

(α) The Gaussian curvature K is defined as above, in terms of areas of normal maps, and is therefore the quotient $\det h_{ik}/\det g_{ik}$. Accordingly, $K \equiv 0$ is equivalent to

$$h_{11}h_{22} - h_{12}^2 \equiv 0. \tag{26}$$

(β) A surface is called a torse if through every point of it there passes at least one line segment lying on the surface and having the property that the tangent plane does not vary along the segment.

(γ) A surface is said to be isometric with a plane at a point (u_0, v_0) if, in some vicinity of (u_0, v_0), there exists a transformation $u^* = u^*(u, v)$, $v^* = v^*(u, v)$ with the properties that it is of class C^1, that its Jacobian does not vanish, and that the element of arc-length on the surface in the (u^*, v^*)-coordinates is Euclidean, $ds^2 = du^{*2} + dv^{*2}$.

It will be shown that, barring an exceptional case, the properties of being a torse, of being isometric with a plane, and having a curvature $K \equiv 0$, are equivalent for surfaces $X(u, v)$ of class C^2.

The first theorem to be proved is

(IV) *Let $X(u, v)$ be a surface of class C^2. The Gaussian curvature is identically 0 if and only if $X(u, v)$ is isometric with a plane at every point.*

The statement to be proved concerning the equivalence of $K \equiv 0$ and of the property of being a torse will involve a qualifying condition:

(V) *Let $X(u, v)$ be a surface of class C^2. The Gaussian curvature is identically 0 if $X(u, v)$ is a torse; conversely, if the Gaussian curvature is identically 0 and if either every point of $X(u, v)$ is a flat point or every point of $X(u, v)$ is a non-flat point, then $X(u, v)$ is a torse.*

By a flat point (u_0, v_0) is meant a point at which the matrix of the second fundamental form is the zero matrix, that is,

$$h_{11} = h_{12} = h_{22} = 0. \tag{27}$$

It will remain an open question whether or not the restriction in the second part of (V), namely, that every point of $X(u, v)$ be of the same type, flat or non-flat, can be dropped.

10. The part of (IV) which states that $K \equiv 0$ implies that $X(u, v)$ is isometric with a plane at every point will be proved first. A standard proof of this implication involves the introduction of geodesic parallel coordinates in a vicinity of a given point, and an application of (5) in those coordinates. It turns out that this method can be adapted to the more general case of the present assumptions. This adaptation involves the replacement of (5) by its integrated form, (7).

By modifying the procedure applied in [3], it can be shown that, in a vicinity of any (u_0, v_0), new parameters u^*, v^* can be introduced as follows: The transformation $u^* = u^*(u, v)$, $v^* = v^*(u, v)$ is of class C^1 and of non-vanishing Jacobian, and the element of arc-length in terms of u^*, v^* is of the form $ds^2 = du^{*2} + G(u^*, v^*)dv^{*2}$, where G (> 0) is continuous and possesses a continuous second partial derivative with respect to u^*. The function G satisfies the Jacobi equation $\partial^2 G^{\frac{1}{2}}/\partial u^{*2} + KG^{\frac{1}{2}} = 0$ and the initial conditions $G(0, v^*) \equiv 1$, $G_{u^*}(0, v^*) \equiv 0$.

Since $K \equiv 0$, these initial conditions imply that $G(u^*, v^*) \equiv 1$. This proves that $X(u, v)$ is isometric with a plane at any point (u_0, v_0).

11. The remaining part of (IV), which states that $K \equiv 0$ if $X(u, v)$ is isometric with a plane at every point, will now be proved. This statement becomes trivial if the surface has a parametrization (near every point) in which the element of arc-length is Euclidean and (5) is valid in that parametrization. But definition (γ) assumes that the parametrization of the surface in which the element of arc-length is Euclidean is only of class C^1; so that not even the integrated form, (7), of (5) is applicable in this coordinate system. This difficulty will be overcome by using a form of (7) which is independent of the coordinate system, namely, the formula of Gauss-Bonnet.

Let (u_0, v_0) be an arbitrary point of the surface $X(u, v)$ of class C^2. Let u^*, v^* denote the parameters described in the definition (γ), whose existence is assumed. Although the surface in the parametric representation (u^*, v^*) might only be of class C^1, it is clear that the (surface) image of any line segment in the (u^*, v^*)-domain is a geodesic (hence, an arc of class C^2), and conversely.

On surfaces of class C^2, the theorem of Gauss-Bonnet is valid ([9], pp. 42-44; cf. [5], p. 136). Hence, if T is the interior of a geodesic triangle, $\alpha + \beta + \gamma + \iint_T K d\sigma = 2\pi$, where $d\sigma$ is the element of surface area and

α, β, γ are the exterior angles of the triangle. Consider small geodesic triangles with the point (u_0, v_0) as a vertex. In the (u^*, v^*)-coordinates, such a geodesic triangle is the image of a triangle in the (u^*, v^*)-plane (in particular, the geodesic triangles exist). Since the (u^*, v^*)-coordinates are Euclidean, the angles α, β, γ are also the exterior angles of the triangle in the (u^*, v^*)-plane. Thus $\alpha + \beta + \gamma = 2\pi$, and so $\iint_T K d\sigma = 0$. Since the domain T is arbitrarily small and since K is continuous, it follows that $K(u_0, v_0) = 0$. This completes the proof of (IV).

12. In order to prove the first part of (V), let $X(u, v)$ be a torse of class C^2. Let (u_0, v_0) be an arbitrary point of the surface and let $u = u(t)$, $v = v(t)$ denote a parametric representation of a line segment which lies on the surface, passes through (u_0, v_0) and has the property that the tangent plane does not vary along this segment. Since the surface is of class C^2, and since the vector product $(X_u, X_v) \neq 0$, the functions $u(t)$, $v(t)$ are of class C^2.

If $N(u, v)$ is the unit surface normal, $dN(u(t), v(t))/dt = 0$. It follows therefore by the formula of Rodrigues that $u = u(t)$, $v = v(t)$ is a line of curvature, and that 0 is a principal curvature along the segment. Hence, the Gaussian curvature, being the product of the principal curvatures, vanishes along the segment; in particular, it vanishes at (u_0, v_0). Since (u_0, v_0) is an arbitrary point, the first part of (V) follows.

13. It remains to prove the second part of (V). It can be supposed that the surface is given in the form $z = z(x, y)$, where z is defined in a vicinity of $(x, y) = (0, 0)$ and is of class C^2. The assumption $K \equiv 0$, being equivalent to (26), reduces to

$$z_{xx} z_{yy} - z_{xy}^2 \equiv 0, \tag{28}$$

since h_{11}, h_{12}, h_{22} are identical with $(1 + z_x^2 + z_y^2)^{-\frac{1}{2}}$ times z_{xx}, z_{xy}, z_{yy}, respectively.

First, if every point is a flat point, that is, if (27) holds for all points (x, y) near $(0, 0)$, then z is a linear function of x and y. Thus the surface is part of a plane, hence a torse.

Next, if $(0, 0)$ is a non-flat point, then at least one of the functions z_{xx}, z_{xy}, z_{yy} does not vanish at $(0, 0)$, and therefore in a vicinity of $(0, 0)$. It can be supposed that $z_{xx}(0, 0) \neq 0$ (otherwise the (x, y)-plane can be subjected to a suitable rotation).

Under the asumption (26), the differential equation of the asymptotic lines, $h_{11}x'^2 + 2h_{12}x'y' + h_{22}y'^2 = 0$, reduces to either of the equivalent equations $h_{11}x' + h_{12}y' = 0$, $h_{12}x' + h_{22}y' = 0$, where the prime denotes differentiation with respect to a parameter which does not occur explicitly. But the last two equations can be written as $z_{xx}x' + z_{xy}y' = 0, z_{yx}x' + z_{yy}y' = 0$, that is, as $(z_x)' = 0$, $(z_y)' = 0$; so that $z_x = \text{const.}$, $z_y = \text{Const.}$ Thus the tangent plane does not vary along an asymptotic curve. Conversely, if, along an arc of class C^1, the function z_x (and/or z_y) is constant, then the arc is an asymptotic arc.

Since $z_{xx} \neq 0$ in a vicinity of $(0, 0)$, it follows from the theorem on implicit functions that, if t is fixed and $|\, t \,|$ is sufficiently small, the equation $z_x(x, y) = z_x(t, 0)$ has a solution $x = x(y, t)$ of class C^1 for $|\, y \,|$, $|\, t \,|$ small. For fixed t, the arc $x = x(y, t)$ is an asymptotic arc, since $z_x(x, y) = \text{const.}$ $(= z_x(t, 0))$. The variables t, y can be introduced as new parameters, since $x(0, t) \equiv t$ implies that $\partial(x, y)/\partial(t, y) = 1$ at $(t, y) = (0, 0)$.

Let $z = z(y; t) = z(x(y, t), y)$. The linear equation

$$a(t)x(y, t) + b(t)y - z(y; t) = c(t), \tag{29}$$

is an identity in y and t if $a(t) = z_x(t, 0)$, $b(t) = z_y(t, 0)$ and $c(t) = a(t)t - z(0; t)$. (Thus, the surface is the envelope of a 1-parameter sheaf of planes, which is of class C^1.) It follows at once that the asymptotic arc $x = x(y, t)$, for t fixed, is a line segment. In fact, differentiation of (29) with respect to t gives

$$a_t(t)x(y, t) + b_t(t)y = c_t(t),$$

since $ax_t - z_t = 0$ is the scalar product of a normal vector and a vector tangent to the (surface) arc $x = x(y, t)$, for y fixed. Since $a_t(t) = z_{xx}(t, 0) \neq 0$, the equations

$$a(t)x + b(t)y - z = c(t) \text{ and } a_t(t)x + b_t(t)y = c_t(t)$$

define planes which intersect along the asymptotic arcs $x = x(y, t)$ for t fixed.

This completes the proof of (V). The proof is an adaptation of one given by Hadamard [2], pp. 398-399.

Part 3. The Fundamental Equations of Curves.

14. The fundamental theorem in the differential geometry of curves states that if $\kappa(s) > 0$ and $\tau(s)$ are, on some s-interval, functions of class C^1 and C^0, respectively, then there exists, on this interval, a curve $X = X(s)$

of class C^3 for which s, $\kappa(s)$ and $\tau(s)$ become the arc-length, the curvature and the torsion, respectively. Furthermore, this $X = X(s)$ is unique, up to translations and rotations of the (x, y, z)-space.

The proof of this theorem involves the Frenet differential equations,

$$U_1' = \kappa U_2, \qquad U_2' = -\kappa U_1 + \tau U_3, \qquad U_3' = -\tau U_2, \tag{30}$$

where the prime denotes differentiation with respect to s and U_1, U_2, U_3 are certain unit vectors in the X-space. Since the linear differential equations (30) possess solutions if the coefficient functions κ, τ are merely continuous, there arises the question as to what remains of the fundamental theorem if it is only assumed that κ (> 0) and τ are continuous. In order to answer this question, it will be necessary to extend the definition of torsion.

Let $X = X(s)$ be an arc of class C^2, of non-vanishing curvature, having its arc-length s as parameter. Let $U_1 = X'$, $U_2 = X''/|X''|$ and $U_3 = (U_1, U_2)$, where the symbol $(\ ,\ )$ is that of vector multiplication, be the unit tangent, the principal normal and the binormal unit vectors, respectively. The arc $X(s)$ will be said to possess a torsion $\tau = \tau(s)$ at a point s if

$$\lim_{\Delta s \to 0} |\, \Delta\theta/\Delta s \,| \tag{31}$$

exists, where $\Delta\theta$ is a continuous determination of the angle between the vectors $U_3(s)$ and $U_3(s + \Delta s)$, such that $\Delta\theta = 0$ for $\Delta s = 0$. When the limit (31) exists it will be called the absolute value of the torsion; so that

$$|\, \tau(s) \,| = \lim_{\Delta s \to 0} |\, (U_3(s), U_3(s + \Delta s))/\Delta s \,| \,.$$

In order to define the sign of $\tau(s)$, note that the definition, $U_3 = (U_1, U_2)$, of the binormal implies that

$$(U_3(s), U_3(s + \Delta s)) = [U_1, U_2, \Delta U_2]U_1 - [U_1, U_2, \Delta U_1]U_2, \tag{32}$$

where $\Delta U_i = U_i(s + \Delta s) - U_i(s)$, and $[A, B, C]$ is the scalar triple product of the vectors A, B, C. Since $\Delta U_1 = (\kappa U_2 + o(1))\Delta s$ as $\Delta s \to 0$, the last term of (32) is $o(|\, \Delta s \,|)$ as $\Delta s \to 0$. Hence the limit (31) exists if and only if $\lim |[U_1, U_2, \Delta U_2/\Delta s]|$ exists (as $\Delta s \to 0$); in which case the two limits are equal. If the last limit exists and is not 0, the limit also exists if the sign of the absolute value is removed. The sign of $\tau(s)$ can be determined from this limit; thus, if (31) exists, put

$$\tau(s) = \lim_{\Delta s \to 0} [U_1, U_2, \Delta U_2/\Delta s]. \tag{33}$$

It is clear that if the torsion exists, in the above sense, at a point, then its

absolute value has the usual geometrical interpretation as the measure of the rate of the twisting of the osculating plane.

Since $U_3 = (U_1, U_2)$ and $U_2 = (U_3, U_1)$, and since U_1 is of class C^1, it is clear that U_2 has a derivative at a given point s if and only if U_3 does. Clearly, a sufficient condition for the existence of the torsion at a point is the existence of U_2' (and/or U_3'); in which case (33) shows that τ is given by the scalar product $U_3 \cdot U_2'$. Furthermore, the Frenet equations (30) hold at such an s. In fact, the first of the equations (30) is the definition of U_2. Since the definition of U_3 implies that U_3' is orthogonal to U_1 and U_3, while $U_3' \cdot U_2 = -U_2' \cdot U_3$, the third of the equations (30) holds. Finally, the second is a consequence of the other two. Conversely, the fundamental theorem can be stated as follows:

(VI) *Let $\kappa(s) > 0$, $\tau(s)$ be continuous functions on an s-interval. Then there exists, on that interval, a curve $X = X(s)$ of class C^2 for which s is arc-length and $\kappa(s)$, $\tau(s)$ are the curvature and the torsion, respectively; in fact, the unit tangent, principal normal, and binormal unit vectors are of class C^1 and satisfy* (30). *The curve $X(s)$ is unique up to translations and rotations of the (x, y, z)-space.*

The proof depends on an adaptation of the proof of the classical theorem. Let $S = (s_{ij})$ denote the three-rowed skew-symmetric matrix in which $s_{12} = \kappa$, $s_{13} = 0$ and $s_{23} = \tau$. If U denotes the matrix whose consecutive columns are U_1, U_2, U_3, then the system of differential equations (30) can be written in the matrix form

$$U' = -US, \qquad (S^* = -S), \tag{34}$$

when the asterisk denotes transposition. It follows from (31) that $U^{*\prime} = SU^*$, hence $(UU^*)' = -USU^* + USU^* = 0$. Thus, $UU^* = \text{const.}$, where const. denotes a constant matrix, is a first integral of (34).

Since the differential equations (30) are linear, the initial conditions can be assigned arbitrarily, and the corresponding solution will exist on the interval on which the coefficient functions are defined. The same holds at every s if the initial conditions are so chosen that, at an initial $s = s_0$, the vectors U_1, U_2, U_3 form a right-hand orthonormal system; that is, in such a way that the matrix U is an orthogonal matrix of determinant $+1$. Of course, such a solution U is uniquely determined up to multiplication (from the left) by arbitrary constant orthogonal matrices of determinant $+1$, since (30) is a linear system. It is clear that, corresponding to such a solution U_1, U_2, U_3

of (30), the indefinite integral $X(s) = \int U_1 ds + C$ defines an arc satisfying the assertions of (VI). The uniqueness of X (up to rotations and translations) is clear from the proof.

15. *Application to Enneper's theorem.* The advantage of (VI) over the usual formulation of the fundamental theorem of curves is due to the fact that (VI) does not assume that the curve is of class C^3 and contains, nevertheless, the theory of torsion. In this regard, the situation becomes particularly revealing after a perusal of the standard form of Enneper's theorem; cf., e. g., [1], p. 147. In fact, all that can then be stated is that if the Gaussian curvature, K, is non-positive on a surface of class C^4, and if Γ is an asymptotic curve of non-vanishing curvature, then the square of the torsion is identical with $-K$ along Γ. Clearly, the C^4-assumption, which is usually not stated, is unnatural indeed. But it cannot be omitted in the usual treatment, since otherwise the asymptotic curve Γ need not be of class C^3, and so the torsion (in the usual sense) of Γ becomes undefined.

On the other hand, if (VI) is used, the C^4-assumption can be reduced to the assumption that the surface be of class C^3. In fact, if the surface is of class C^3, the asymptotic arc Γ is of class C^2. Since Γ is of non-vanishing curvature, its binormal, U_3, is the surface normal, N, which is of class C^2; hence, (VI) is applicable to Γ.

Despite the fact that, for the existence of the torsion, it is sufficient that U_3 be only of class C^1, the theorem of Enneper becomes meaningless if it is only assumed that the surface is of class C^2. For an asymptotic arc might then be of class C^1; so that the principal normal U_2, hence the binormal U_3, need not be defined. In order to see this, consider the surface $z = g(x) + xy + \frac{1}{2}y^2$, where $g(x)$ is of class C^2. The equations for the asymptotic lines reduce to $g''x'^2 + 2x'y' + y'^2 = 0$, that is, $dy/dx = -1 \pm (1 - g'')^{\frac{1}{2}}$. Clearly, the Gaussian curvature is negative if and only if $1 - g'' > 0$. Let $h(x) = -1 \pm (1 - g'')^{\frac{1}{2}}$; so that

$$g'' = -2h - h^2.$$

Let $h(x)$ be an arbitrary continuous function which vanishes at $x = 0$ and is not of class C^1. Finally, let $g(x)$ be defined by the last formula line (while $g(0)$ and $g'(0)$ are arbitrary). Then the surface, z, is of class C^2 and has a negative curvature, for $|x|$ sufficiently small. It is clear that the asymptotic arcs will not be of class C^2 if $h(x)$ is suitably chosen.

By a suitable choice of an $h(x)$ of class C^1 in the above considerations, it is also seen that, on surfaces of class C^3, the set of those points of on

asymptotic arc at which its curvature is 0 can be, for example, a nowhere dense, perfect set.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] A. Duschek and W. Mayer, *Lehrbuch der Differentialgeometrie*, vol. 1, Leipzig and Berlin, 1930.

[2] J. Hadamard, *Cours d'analyse*, vol. 1, Paris, 1927.

[3] P. Hartman, "On the local uniqueness of geodesics," *American Journal of Mathematics*, vol. 72 (1950), pp. 723-730.

[4] ——— and A. Wintner, "On the embedding problem in differential geometry," *ibid.*, vol. 72 (1950), pp. 553-564.

[5] E. R. van Kampen, "The theorems of Gauss-Bonnet and Stokes," *ibid.*, vol. 60 (1938), pp. 129-138.

[6] H. Lebesgue, "Intégrale, longueur, aire," *Annali di Matematica Pura ed Applicata*, ser. 3, vol. 7 (1902), pp. 231-359.

[7] W. Nikiliborc, "Sur les équations linéaires aux différentielles totales," *Studia Mathematica*, vol. 1 (1929), pp. 41-49.

[8] F. Schur, "Ueber den analytischen Charakter der eine endliche continuierliche Transformationsgruppe darstellenden Functionen," *Mathematische Annalen*, vol. 41 (1893), pp. 509-538.

[9] H. Weyl, "Ueber die Bestimmung einer geschlossenen konvexen Fläche durch ihr Linienelement," *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, vol. 61 (1916), pp. 40-72.

LAMÉ COORDINATES IN HILBERT SPACE.*

By PHILIP HARTMAN and AUREL WINTNER.

---

1. In an n-dimensional unitary space, let $\xi = (\xi_1, \xi_2, \cdots, \xi_n)$ be a point not on a parameter plane (so that $\xi_k \neq 0$ for $k = 1, \cdots, n$), and let $\alpha_1 < \alpha_2 < \cdots < \alpha_n$. Then the equation

$$\sum_{k=1}^{n} |\xi_k|^2/(\alpha_k - \lambda) = 1 \tag{1}$$

has n distinct real roots, say, $\lambda_1 < \lambda_2 < \cdots < \lambda_n$, which separate, and are separated by, the numbers $\alpha_1, \cdots, \alpha_n$. This is the classical definition of the Lamé coordinates $\lambda_1, \cdots, \lambda_n$ of ξ. The following considerations will be made possible by the fact that it is possible to introduce the Lamé coordinates by quite another, though equivalent, definition, as follows:

Let (x, y) denote the scalar product, $x_1\bar{y}_1 + \cdots + x_n\bar{y}_n$, of x and y, and let $\phi = c^{-\frac{1}{2}}\xi$, where $c = (\xi, \xi)$ is the squared length of ξ. If H denotes the diagonal matrix having the diagonal elements $\alpha_1, \alpha_2, \cdots, \alpha_n$, and if I is the unit matrix, then (1) can be written in the form $((H - \lambda I)^{-1}\phi, \phi) = -1/c$. Let E denote the matrix of the linear transformation

$$Ex = (x, \phi)\phi; \tag{2}$$

so that E is a one-dimensional projection. It turns out that the eigenvalues of the Hermitian matrix $H + cE$ are the Lamé coordinates, $\lambda_1, \cdots, \lambda_n$, of ξ.

In this characterization of the Lamé coordinates of ξ, the restriction that all coordinates ξ_k of ξ be different from 0 can be omitted. This generalization merely allows the possibility that some of the Lamé coordinates $\lambda_1, \cdots, \lambda_n$ of ξ and some of the numbers $\alpha_1, \cdots, \alpha_n$ are equal. (Of course, there still remains the difficulty that $(\xi_1, \cdots, \xi_n)$ and $(e^{i\theta_1}\xi_1, \cdots, e^{i\theta_n}\xi_n)$, for arbitrary real $\theta_1, \theta_2, \cdots, \theta_n$, have the same Lamé coordinates.) It is clear that, in this definition, H need not be a diagonal matrix with distinct roots but can be an arbitrary Hermitian matrix. In the latter case, the number α occurs at least $k - 1$ times among the Lamé coordinates of every element ξ, if the characteristic number $\lambda = \alpha$ of H is of multiplicity k.

2. The object of this note is to carry over these ideas to Hilbert's space. Let H denote an arbitrary Hermitian matrix, bounded in Hilbert's space,

---

* Received April 5, 1950.

ϕ a vector of length 1 in this space, and E the corresponding projection (2). Associated with H and ϕ is a bounded Hermitian matrix, say K, which is the projection of H on the hyperspace orthogonal to ϕ. Thus, $K = (I-E)H(I-E)$, where I is the unit matrix. For example, if the coordinates of ϕ are $(1, 0, 0, \cdots)$, then K is obtained by removing from H the first row and first column. The matrix K is considered as operating on the hyperplane orthogonal to ϕ. Clearly,

$$K = (I-E)(H+cE)(I-E), \text{ where } -\infty < c < \infty.$$

It was recently shown ([3], p. 867) that the spectra of K and H separate each other; in the sense that if λ_1 and λ_2, where $\lambda_1 < \lambda_2$, are two points of the spectrum of K (or H), then the interval $\lambda_1 \leqq \lambda \leqq \lambda_2$ contains at least one point of the spectrum of H (or K). The same holds in the limiting case $\lambda_1 = \lambda_2$, if $\lambda = \lambda_1 = \lambda_2$ is in the point spectrum of K (or H) with a multiplicity $m > 1$; in which case λ is in the point spectrum of H (or K) with at least the multiplicity $m-1$.

These separation theorems imply that the cluster points of the spectra of K and H are identical. This corollary is contained in a theorem of Weyl [9], p. 384, also, since the difference $K-H$ is completely continuous.

The above comments make clear the content of the following theorems:

(I) *Let the spectrum of K contain the points λ_1, λ_2 but no point of the interval $\lambda_1 < \lambda < \lambda_2$. Then, to every λ on this interval, there belongs a unique real $c = c(\lambda)$, corresponding to which λ is in the spectrum of $H + cE$. The function $c = c(\lambda)$, given by*

$$c(\lambda) = -1/((H-\lambda I)^{-1}\phi, \phi), \tag{3}$$

is regular-analytic and increasing for $\lambda_1 < \lambda < \lambda_2$.

That (3) is meaningful, that is, that $((H-\lambda I)^{-1}\phi, \phi) \neq 0$, will be shown in the proof of (I). It is understood that $c(\lambda) = 0$ if λ is in the spectrum of H.

(II) *Let the conditions of* (I) *hold and let $\lambda = \lambda_1$ be an isolated point of the point spectrum of K, hence of $H + cE$. Let m, $m(c)$ denote the multiplicity of λ_1 in the point spectrum of K, $H + cE$, respectively, so that $0 \leqq m-1 \leqq m(c) \leqq m+1$. Finally, let $c_1 = \lim c(\lambda)$ as $\lambda \to \lambda_1 + 0$, where $-\infty \leqq c_1 < \infty$. Then $c_1 = -\infty$ holds if and only if $m(c) = m-1$ is an identity for $-\infty < c < \infty$.*

(II bis) *If $c_1 > -\infty$, then $m(c) = m$ or $m(c) = m+1$ according as $c \neq c_1$ or $c = c_1$.*

3. Proof of (I). If λ is on the interval (λ_1, λ_2), then there is at most one c corresponding to which λ is in the spectrum of $H + cE$. For suppose, if possible, that there exist two distinct values c_1, c_2 with the property that λ is in the spectra of both $H + c_1E$ and $H + c_2E$. Let x_1, x_2 denote normalized eigenfunctions satisfying

$$(H + cE)x = \lambda x \tag{4}$$

for $c = c_1$, c_2, respectively. The elements x_1, x_2 are either linearly independent or dependent. In the first case, let x denote a linear combination of x_1, x_2 of length 1 and orthogonal to ϕ, that is, $|x| = 1$ and $Ex = 0$. It then follows from (4) that $Hx = \lambda x$. In the second case, it can be supposed that $x = x_1 = x_2$. Then (4) shows that $(c_1 - c_2)Ex = 0$; hence $Ex = 0$, since $c_1 \neq c_2$. Thus, in either case, there exists an element x, of length 1, satisfying $Hx = \lambda x$ and $Ex = 0$. But the definition of K implies that

$$Kx = \lambda x, \tag{5}$$

which contradicts the assumption that no point of the interval (λ_1, λ_2) is in the spectrum of K.

The separation theorem, quoted before (I), shows that the interval (λ_1, λ_2) contains at most one point of the spectrum of H. If such a point exists, denote it by λ_0. If $\lambda \neq \lambda_0$, let $x = (H - \lambda I)^{-1}\phi$. Then x exists and is not 0. Furthermore, $(x, \phi) = ((H - \lambda I)^{-1}\phi, \phi) \neq 0$. For otherwise $(I - E)x = x$ holds, hence $(K - \lambda(I - E))x$ becomes $(I - E)(H - \lambda I)x$, which is $(I - E)\phi$ and therefore 0. But this contradicts the assumption that λ is not in the spectrum of K. Consequently, (3) is meaningful if $\lambda = \lambda_0$. The regular-analyticity and monotony of $c(\lambda)$ at $\lambda \neq \lambda_0$ follow from the identity

$$((H - \lambda I)^{-1}\phi, \phi) = \int_{-\infty}^{+\infty} (\mu - \lambda)^{-1} d\rho(\mu), \tag{6}$$

where $\rho(\mu)$ is the non-decreasing function on $-\infty < \mu < \infty$ defined by $\rho(\mu) = (E(\mu)\phi, \phi)$ in terms of the resolution of the identity $E(\mu)$ belonging to H.

The definitions of x, E and $K = (H - \lambda I)^{-1}\phi$ show that $(H - \lambda I + cE)x = \phi + cEx = \phi + c((H - \lambda I)^{-1}\phi, \phi)\phi$. If c is given by (3), it is seen that (4) holds; so that λ is in the point spectrum of $H + cE$.

It remains to show that if λ_0 exists, then $c(\lambda)$ is regular-analytic and increasing at $\lambda = \lambda_0$. (Of course, $c(\lambda_0) = 0$.) To this end, let $c_0 \neq 0$ be a real number; so that λ_0 (and hence any λ near λ_0) is not in the spectrum of $H_0 = H + c_0E$. Since the relation between K and $H + cE$ is independent

of c, it follows that if λ is on the interval (λ_1, λ_2), then λ is in the spectrum of $H_0 + cE$ if and only if $c = c(\lambda) - c_0$. Thus, for λ near λ_0,

$$(7) \qquad c(\lambda) - c_0 = -1/((H_0 - \lambda I)^{-1}\phi, \phi).$$

This completes the proof of (I).

4. Proof of (II). It will first be shown that if $m(c) > m - 1$ for one value of c, say for $c = c_0$, then $m(c) > m - 1$ for all values of c.

Let $x_1, \cdots, x_m$ denote m ortho-normalized solutions of (4), where $c = c_0$ and $\lambda = \lambda_1$. If $(x_j, \phi) = 0$ for $j = 1, \cdots, m$, then $x_1, \cdots, x_m$ are solutions of (4), with $\lambda = \lambda_1$, implying that $m(c) \geqq m$ holds for all values of c. It can, therefore, be supposed that $(x_1, \phi) \neq 0$. It can also be supposed that $(x_j, \phi) = 0$ for $j = 2, \cdots, m$. For otherwise x_j, where $j = 2, \cdots, m$, can be replaced by a suitable linear combination of x_j and x_1 (then x_1 becomes replaced by an element orthogonal to $x_2, \cdots, x_m$). Consequently, the elements $x_2, \cdots, x_m$ are also solutions of (5), with $\lambda = \lambda_1$. Since $\lambda = \lambda_1$ has a multiplicity m in the point spectrum of K, there exists a normalized element $x = y$ satisfying (5) with $\lambda = \lambda_1$, $(y, \phi) = 0$ and $(y, x_j) = 0$ for $j = 2, \cdots, m$.

On the other hand, the definition of K and the relations $(y, \phi) = 0$ and (5), with $\lambda = \lambda_1$, imply that there exists a constant k satisfying $(H - \lambda I)y = k\phi$. It can be supposed that $k \neq 0$, for otherwise $y, x_2, \cdots, x_m$ are m ortho-normalized solutions of (4), with $\lambda = \lambda_1$, for every c, and the assertion is proved.

If $k \neq 0$, it is easily verified that $\alpha y + x_1$ is a solution of (4), where $\lambda = \lambda_1$ and $\alpha = (c_0 - c)(x_1, \phi)/k$. Clearly, $\alpha y + x_1$ is orthogonal to $x_2, \cdots, x_m$; and $\alpha y + x_1 \neq 0$, since $(\alpha y + x_1, \phi) = (x_1, \phi) \neq 0$. Hence (4), with $\lambda = \lambda_1$, has at least m linearly independent solutions for every c.

5. The assertions of (II) can now be proved as follows:

First, it will be shown that if $m(c) = m - 1$ holds for all values of c, then $c_1 = -\infty$.

Suppose, if possible, that $c_1 > -\infty$. Let $y, x_2, \cdots, x_m$ denote m ortho-normalized solutions of (5), with $\lambda = \lambda_1$, chosen so as to satisfy $(y, \phi) = (x_j, \phi) = 0$ and $(H - \lambda_1 I)x_j = 0$ for $j = 2, \cdots, m$. Let $x(\lambda)$ denote a normalized solution of (4), for $c = c(\lambda)$ and λ $(> \lambda_1)$ near λ_1. There exists a sequence of λ-values $\lambda^1, \lambda^2, \cdots$ satisfying $\lambda^n > \lambda^{n+1} \to \lambda_1$ as $n \to \infty$, and $x_1 = \lim x(\lambda^n)$ exists in the weak sense. Since, as $n \to \infty$,

$$(8) \qquad (H - \lambda_1 I + c_1 E)x(\lambda^n) = (\lambda^n - \lambda_1)x(\lambda^n) + (c_1 - c(\lambda))x(\lambda^n)$$

tends to 0 in the strong sense, the limit x_1 is not 0. For, if it were, it would follow from Weyl's criterion ([9], p. 378) that $\lambda = \lambda_1$ is a cluster point of

the spectrum of $H + c_1E$, and hence of K. Consequently, $x = x_1 \neq 0$ satisfies (4) for $c = c_1$ and $\lambda = \lambda_1$. Also, $(x_1, x_j) = 0$ for $j = 2, \cdots, m$. In fact, $(x(\lambda^n), x_j) = 0$ for $j = 2, \cdots, m$ and $n = 1, 2, \cdots$, since $x(\lambda^n)$ and x_j satisfy (4) for $c = c(\lambda^n)$, but for the distinct λ-values, $\lambda = \lambda^n$ and λ_1, respectively. Since this implies that $m(c_1) \geqq m$, it contradicts the assumption that $m(c) = m - 1$ for every c.

In order to prove the converse, let $c_1 = -\infty$. Suppose, if possible, that $m(c) \geqq m$ for all values of c. The assumption that $\lambda = \lambda_1$ is an isolated point of the spectrum of H shows that the monotone function $\rho(\mu)$ in (6) is constant near $\mu = \lambda_1$, except for a possible jump at $\mu = \lambda_1$.

It follows from (3) that $c_1 = -\infty$ implies that $((H - \lambda I)^{-1}\phi, \phi)$ has a finite limit, viz. 0, as $\lambda \to \lambda_1 + 0$. Hence, the monotone function $\rho(\mu)$ has no jump at $\mu = \lambda_1$. Thus, ϕ is orthogonal to the linear space spanned by the solutions of $(H - \lambda_1 I)x = 0$. Let $x_1, \cdots, x_m$ denote m ortho-normalized solutions of this equation. Then $x_1, \cdots, x_m$ are solutions of (5) with $\lambda = \lambda_1$.

Since $\rho(\mu)$ is constant near $\mu = \lambda_1$, it follows that there exists an x, orthogonal to $x_1, \cdots, x_m$, satisfying the equation $(H - \lambda_1 I)x = \phi$. In fact, x is given by

$$x = \{\int_{-\infty}^{+\infty} (\mu - \lambda_1)^{-1} dE(\mu)\}\phi \tag{9}$$

(cf., e. g., the Appendix below). Clearly, x satisfies $(x, \phi) = 0$, since $((H - \lambda I)^{-1}\phi, \phi) \to 0$ as $\lambda \to \lambda_1 + 0$. Furthermore, it is seen from the definition of K that (5) holds for $\lambda = \lambda_1$. It is also clear that $x_1, \cdots, x_m$ and x are linearly independent, since $(H - \lambda_1 I)x_j = 0$ for $j = 1, \cdots, m$, while $(H - \lambda_1 I)x = \phi$. But this contradicts the assumption that the multiplicity of $\lambda = \lambda_1$ in the point spectrum of K is only m. Thus the proof of (II) is complete.

6. Proof of (II bis). Let $c_1 > -\infty$; so that $m(c) \geqq m$ for all values c. Let $x_1(c), x_2, \cdots, x_m$ denote m ortho-normal solutions of (4), with $\lambda = \lambda_1$, satisfying $(x_j, \phi) = 0$ for $j = 2, \cdots, m$. If $\lambda\ (> \lambda_1)$ is near λ_1, let $x(\lambda)$ denote a normalized solution of (4), where $c = c(\lambda)$. Then $x(\lambda)$, $x_1(c(\lambda))$, $x_2, \cdots, x_2$ form an ortho-normal set. Let $y(\lambda) = \alpha x(\lambda) + \beta x_1(c(\lambda))$, where $\alpha = \alpha(\lambda)$ and $\beta = \beta(\lambda)$, denote a normalized linear combination of $x(\lambda)$ and $x_1(c(\lambda))$ satisfying $(y(\lambda), \phi) = 0$. There exists a sequence of λ-values satisfying $\lambda^n > \lambda^{n+1} \to \lambda_1$ as $n \to \infty$, and $y_1 = \lim y(\lambda^n)$ exists in the weak sense. Since $|\alpha|^2 + |\beta|^2 = 1$, it follows that

$$(H-\lambda_1 I)y(\lambda^n) = (H-\lambda_1 I + c(\lambda^n)E)y(\lambda^n) = \alpha(\lambda^n-\lambda_1)x(\lambda^n)$$

tends to 0 in the strong sense, as $n \to \infty$. Consequently, the limit y_1 of $y(\lambda^n)$ is not 0. For, if it were, it would follow from Weyl's criterion used above, that $\lambda = \lambda_1$ is in the essential spectrum of H.

Let $y = y_1/|\,y_1\,|$; so that $y, x_2, \cdots, x_m$ form an ortho-normal set, each member of which is orthogonal to ϕ and satisfies (4), with $\lambda = \lambda_1$, for every value of c. By repeating the above argument involving (8), it follows that there exists an element x_1 orthogonal to $y, x_2, \cdots, x_m$ and satisfying (4), with $\lambda = \lambda_1$ and $c = c_1$. Hence, $m(c_1) = m + 1$.

It remains to show that there cannot exist another value of c, say c_2, distinct from c_1 and satisfying $m(c_2) = m + 1$.

Suppose the contrary. Let $x(c_1), y, x_2, \cdots, x_m$ and $x(c_2), y, x_2, \cdots, x_m$ denote ortho-normal sets of solutions of (4), with $\lambda = \lambda_1$, belonging to $c = c_1$ and c_2, respectively. Whether or not $x(c_1)$ and $x(c_2)$ are linearly independent, they have a normalized linear combination, say x^1, which is orthogonal to ϕ; cf. the proof of (I), above. But then $x^1, y, x_2, \cdots, x_m$ are solutions of (5), with $\lambda = \lambda_1$. Since this is a contradiction, the proof of (II bis) is complete.

7. The signatures of the Lamé quadrics in Hilbert space. In the notation of §1, let H be an n-rowed diagonal matrix with $\alpha_1 < \cdots < \alpha_n$ as diagonal elements, and let $\lambda_1 < \cdots < \lambda_n$ be the Lamé coordinates of $\xi = (\xi_1, \cdots, \xi_n)$, where $\xi_1 \neq 0, \cdots, \xi_n \neq 0$. Then, if $k = 1, \cdots, n$, the number of the negative characteristic numbers of $H - \lambda_k I$ is $k - 1$. Correspondingly, for $\lambda = \lambda_1, \cdots, \lambda_n$ the quadric (1) attains each of the n types possible for a (real) central quadric (namely, the $n-1$ types of hyperboloids and the real ellipsoid). In what follows, this fact concerning consecutive signatures will be extended in the direction of (I)-(II bis) to the case of Hilbert's space.

Let H be a bounded self-adjoint matrix having only simple points (if any) in its point spectrum, and ϕ a vector of length 1 in Hilbert space with the property that no characteristic vector of H is orthogonal to ϕ. Thus, $x = 0$ is implied by $(x, \phi) = 0$ and $Hx = \lambda x$. These assumptions on H and ϕ imply that, for a fixed value of c,

(*) $\qquad H + cE$ does not have a multiple eigenvalue.

For if (*) is false, then there exists a $\lambda = \lambda_c$ corresponding to which (4) has two linearly independent solutions, say x_1 and x_2. Hence, a suitable linear combination x of x_1 and x_2 will be such as to satisfy both $x \neq 0$ and $(x, \phi) = 0$. It follows therefore from (2) that $Ex = 0$, and from (4) that

$Hx = \lambda x$. But the existence of such an x contradicts the assumptions concerning H and ϕ. This proves (*).

Let $c > 0$ and $\xi = c\phi$, so that the "Lamé coordinates of ξ" correspond to the points of the spectrum of $H + cE$. In view of the remarks preceding the statement of (I), the set of the cluster points of the spectrum of $H + cE$, for a fixed c, is identical with the set of cluster points of the spectrum of H. Let λ' and λ'' respectively denote the least and the greatest cluster point of the spectrum of H. Then (I) implies that the (empty, finite or denumerable) set of points $\lambda_1 < \lambda_2 < \cdots$ of the spectrum of $H + cE$ ("Lamé coordinates of ξ") which are less than λ' are determined by the equation $((H - \lambda I)^{-1}\phi, \phi) = -1/c$. Clearly, $-((H - \lambda I)^{-1}\phi, \phi)^{-1} \to +0$ as $\lambda \to -\infty$. Also, if λ_0 $(< \lambda')$ is a point of the spectrum of H, then it follows from (I) that $-((H - \lambda I)^{-1}\phi, \phi)^{-1}$ tends to ∞ or $-\infty$ according as $\lambda \to \lambda_0 - 0$ or $\lambda \to \lambda_0 + 0$.

These facts, when combined with (*), show that the number of negative points in the spectrum of $H - \lambda_k I$ is $k - 1$. A corresponding statement applies to those "Lamé coordinates of ξ" which exceed λ''. Finally, it is clear from the definition of λ', λ'' that if $\lambda' < \lambda''$, and if λ satisfies $\lambda' < \lambda < \lambda''$ and is in the spectrum of $H + cE$, then there is an infinite number of positive and of negative points in the spectrum $H - \lambda I$.

8. An application to differential operators. Let $p(t)$ be a positive, and $q(t)$ a real-valued, continuous function for $0 \leqq t < \infty$, and let the differential equation

$$(9) \qquad (px')' + (\lambda + q)x = 0$$

be of *Grenzpunkt* type. Then (9) and a homogeneous boundary condition, say

$$(10) \qquad x(0)\sin\alpha - p(0)x'(0)\cos\alpha = 0,$$

determine a self-adjoint eigenvalue problem with a real spectrum $S(\alpha)$. Let $T(x, \alpha) = -(px')' - qx$ denote the self-adjoint operator associated with (10); so that the spectrum of $T(x, \alpha)$ is $S(\alpha)$. The set of cluster points S' of $S(\alpha)$ is independent of α; cf. [10], p. 251. Suppose that some real λ-value, say $\lambda = 0$, is not in S'. Then there is a value α_0 of α (unique mod π), corresponding to which $\lambda = 0$ belongs to $S(\alpha_0)$; cf. [1].

Let $x = \phi(t)$ denote a normalized solution of (9) and (10) for $\lambda = 0$ and $\alpha = \alpha_0$; so that $T(\phi, \alpha_0) = 0$. On the sub-domain of $T(x, \alpha_0)$ orthogonal to ϕ, the operator $T(x, \alpha_0)$ possesses a bounded inverse K, since $\lambda = 0$ is an isolated point of $S(\alpha_0)$. If $\alpha \not\equiv \alpha_0 \pmod{\pi}$, then $T(x, \alpha)$ has a bounded inverse $H(\alpha)$, which is an integral operator; cf. [2].

Let E be the projection of the Hilbert space $L^2(0,\infty)$ on the 1-dimensional linear subspace containing ϕ and let $H = H(\alpha_0 + \frac{1}{2}\pi)$. It is easily verified, from the form of the Green kernel belonging to $H(\alpha)$, that $H(\alpha) = H + cE$, where $c = c(\alpha)$ is a constant (independent of α) multiple of $\cot(\alpha_0 - \alpha)$ for $\alpha \not\equiv \alpha_0 \pmod{\pi}$. Furthermore, $K = (I - E)H(\alpha)(I - E)$.

These facts, when combined with theorems (I) and (II), imply the results of [4] concerning the separation of the spectral sets $S(\alpha)$, $S(\beta)$ for $\alpha \not\equiv \beta \pmod{\pi}$.

9. Another illustration: Mathieu functions. For a fixed $q > 0$, let $\lambda_0(q) < \lambda_1(q) < \cdots$ denote the λ-values corresponding to which Mathieu's differential equation,

$$x'' + (\lambda + 16q\cos^2 t)x = 0, \tag{11}$$

has a solution ($\not\equiv 0$) of period π or 2π. It will be shown that the results of Sections **2** and **7** imply that

(i) if $q > 0$ and λ are fixed, then (11) cannot possess two linearly independent solutions of period π or 2π; and that

(ii) according as n in $\lambda = \lambda_n(q)$ is of the form $n = 4k$, $4k+3$, $4k+2$, $4k+1$, the differential equation (11) possesses a solution of the form

$$(12_1)\quad x(t) = \tfrac{1}{2}x + \sum_{m=1}^{\infty} x_m \cos 2mt; \qquad (12_2)\quad x(t) = \sum_{m=1}^{\infty} x_m \sin 2mt,$$

$$(13_1)\quad x(t) = \sum_{m=0}^{\infty} x_m \cos(2m+1)t; \qquad (13_2)\quad x(t) = \sum_{m=0}^{\infty} x_m \sin(2m+1)t,$$

respectively.

(i) is a theorem of Ince [5]; while (ii) can be derived as a consequence of an asymptotic formula obtained by Ince [6] for $\lambda_k(q)$ as $q \to \infty$ (when k is fixed). Actually, it is well known that if $\lambda = \lambda_0(q)$, then (11) has a solution of the form (12_1); that if λ is $\lambda_{4n+1}(q)$ or $\lambda_{4n+2}(q)$, then (11) has a solution of period 2π (but not of period π), so that the corresponding periodic solution is of one of the forms (13_1) or (13_2); that if λ is $\lambda_{4n+3}(q)$ or $\lambda_{4n}(q)$, then the periodic solution of (11) is of one of the forms (12_1) or (12_2). The point of the assertion (ii) is that $\lambda = \lambda_{4n+1}(q)$ corresponds to a periodic solution of the form (13_1), and $\lambda = \lambda_{4n+2}(q)$ to one of the form (13_2), etc.

In the following proof of (i) and (ii), it will be necessary to use the results of Sections **2** and **7** for the case of a self-adjoint matrix H which is not bounded. The nature of the above proofs makes it clear that this extension is obvious.

Since the factor $\cos 2x$ in (11) is even, any real periodic solution can be assumed to be either even or odd. A comparison of coefficients shows that if (11) has a periodic solution, it can be supposed to have one of the forms (12_1), (12_2), (13_1), (13_2).

Let $H = (h_{nm})$, where $m, n = 0, 1, \cdots$, be the Jacobi matrix in which h_{nm} is $4n^2$, $-8q$ or 0 according as $n = m$, $n - m = \pm 1$ or $|n - m| > 1$. Let ϕ be the vector having the coordinates $(1, 0, 0, \cdots)$ in Hilbert space. Then (2) shows that E is the matrix (e_{nm}) in which every e_{nm}, except $e_{00} = 1$, is 0. Correspondingly, K is the matrix which results if the first row and first column of H are removed; that is, $K = (h_{mn})$, where $m, n = 1, 2, \cdots$. A comparison of coefficients shows that (11) has a solution of the form (12_1) if and only if there exists a point $x = (x_0, x_1, \cdots) \neq 0$ of Hilbert space satisfying

$$(H - \lambda I + \tfrac{1}{2}\lambda E)x = 0; \tag{14}$$

and that (11) has a solution of the form (12_2) if and only if there exists a point $x = (x_1, x_2, \cdots) \neq 0$ in the Hilbert hyperplane $x_0 = 0$ satisfying

$$(K - \lambda(I - E))x = 0. \tag{15}$$

The considerations of Section 7 (cf. (*)) show that $H + cE$ and K cannot have a common eigenvalue if no eigenvector $x = (x_0, x_1, \cdots)$ of H is orthogonal to ϕ. But the form of the equations $(H - \lambda I)x = 0$ shows that if $(x, \phi) = 0$, that is, if $x_0 = 0$, then $x_1 = x_2 = \cdots = 0$. Hence, there exists no value of λ corresponding to which both (14) and (15) have non-trivial solutions

It follows from (I) and (II)-(II bis), Section 2, that the values of λ corresponding to which (14) has non-trivial solutions, separate and are separated by those for which (15) has non-trivial solutions. Hence, the periodic solutions of (11) are of the form (12_1) or (12_2) according as n in $\lambda = \lambda_n(q)$ is of the form $n = 4k$ or $n = 4k + 3$.

In order to consider solutions of the form (13_1) or (13_2), let $H = (h_{mn})$, where $m, n = 0, 1, \cdots$, be the Jacobi matrix defined as follows: If $n \neq 0$ and $m \neq 0$, then h_{mn} is $(2n + 1)^2$, $-8q$ or 0 according as $n = m$, $n = m \pm 1$ or $|n - m| > 1$; while if either $n = 0$ or $m = 0$, then h_{mn} is $1 - 8q$, $-8q$ or 0 according as $n = m$, $n = m \pm 1$ or $|n - m| > 1$. As before, let $\phi = (1, 0, 0, \cdots)$. A comparison of coefficients shows that (11) has a solution of the form (13_1) if and only if there exists a point $x = (x_0, x_1, x_2, \cdots) \neq 0$ in Hilbert space satisfying

$$(H-\lambda I)x=0; \tag{16}$$

and that (11) has a solution of the form (13_2) if and only if there exists in the same space a non-trivial solution x of

$$(H-\lambda I+16qE)x=0. \tag{17}$$

As above, (16) and (17) cannot have non-trivial solutions for the same value of λ. It follows that the eigenvalues of H and $H+16qE$ separate each other; and that, since $c=16q>0$, the least eigenvalue of $H+16qE$ is greater than that of H. Hence, the periodic solutions of (11) are of the form (13_1) or (13_2) according as n in $\lambda=\lambda_n(q)$ is of the form $n=4k+1$ or $n=4k+2$.

Appendix.

Let A be a completely continuous matrix; $\lambda_1, \lambda_2, \cdots$ its eigenvalues; $\phi_1, \phi_2, \cdots$ a corresponding set of eigenfunctions; y a point of Hilbert's space; finally,

$$y \sim \sum_{k=1}^{\infty} a_k \phi_k \tag{1}$$

the "Fourier series" of y. Picard [7] has shown that the infinite system of linear equations

$$Ax=y \tag{2}$$

has a solution (in Hilbert space) if and only if

$$\sum_{k=1}^{\infty} |a_k|^2/|\lambda_k|^2 < \infty; \tag{3}$$

in which case

$$x \sim \sum_{k=1}^{\infty} a_k \phi_k/\lambda_k \tag{4}$$

is a solution of (2), the (unique) solution of least length. In (3), it is understood that if $\lambda_k=0$ for some k, then $|a_k|^2/|\lambda_k|^2$ denotes ∞ or 0 according as $a_k \neq 0$ or $a_k=0$.

The object of this Appendix is to point out the corresponding criterion for the solvability of (2) when A is not completely continuous, but just bounded, say.

Suppose first that A is self-adjoint and let $E(\lambda)$, where $-\infty<\lambda<\infty$,

be its resolution of the identity. The equations (2) are solvable if and only if

$$\int_{-\infty}^{+\infty} \lambda^{-2} d \mid E(\lambda) y \mid^2 < \infty ; \tag{5}$$

in which case

$$x = \{ \int_{-\infty}^{+\infty} \lambda^{-1} dE(\lambda) \} y \tag{6}$$

is a solution of (2), the solution of least length. This follows from the fact that a necessary condition for the solvability of (2) is that y be on the manifold M orthogonal to the solutions of the homogeneous equations $Ax = 0$. On a subset of M, the matrix A has a unique inverse A^{-1}, not necessarily bounded, given by

$$A^{-1} = \int_{-\infty}^{+\infty} \lambda^{-1} dE'(\lambda),$$

where $E'(\lambda) = E(\lambda)$ or $E(\lambda) - E(0)$ according as $\lambda < 0$ or $\lambda \geqq 0$. Finally, (5) is necessary and sufficient in order that y be in the domain of A^{-1}. Then (6) is $A^{-1}y$. Since (6) is orthogonal to all solutions of $Ax = 0$, it follows that (6) is the solution of (2) of least length.

If A is not self-adjoint, it is sufficient to apply the well-known device of replacing (2) by the equivalent system of equations

$$\begin{pmatrix} 0 & A \\ A^* & 0 \end{pmatrix} \begin{pmatrix} 0 \\ x \end{pmatrix} = \begin{pmatrix} y \\ 0 \end{pmatrix},$$

in which the operator represented by the 2-rowed matrix is self-adjoint (A^* denotes the adjoint of A). What thus results, in case of solvability, is again the solution of least length. Correspondingly, the above result can be thought of as a spectral formulation of the criteria of E. Schmidt [8], which he expresses in terms of limits of Gram determinants or equivalent "elementary" procedures.

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] P. Hartman and A. Wintner, "An oscillation theorem for continuous spectra," *Proceedings of the National Academy of Sciences*, vol. 33 (1947), pp. 376-379.

[2] ——— and A. Wintner, "On the orientation of unilateral spectra," *American Journal of Mathematics*, vol. 70 (1948), pp. 309-316.

[3] ——— and A. Wintner, "Separation theorems for bounded Hermitian forms," *ibid.*, vol. 71 (1949), pp. 865-878.

[4] ——— and A. Wintner, "A separation theorem for continuous spectra," *ibid.*, vol. 71 (1949), pp. 650-663.

[5] E. L. Ince, "A proof of the impossibility of the coexistence of two Mathieu functions," *Proceedings of the Cambridge Philosophical Society*, vol. 21 (1922), pp. 117-120.

[6] ———, "The Mathieu equation with numerically large parameters," *Journal of the London Mathematical Society*, vol. 2 (1927), pp. 46-50.

[7] É. Picard, "Sur un théorème général relatif aux équations intégrales de première espèce," *Rendiconti del Circolo Matematico di Palermo*, vol. 29 (1910), pp. 79-96, § 1 (reprinted in Picard's *Selecta*, Paris (1928), pp. 206-211).

[8] E. Schmidt, "Ueber die Auflösung linearer Gleichungen mit unendlich vielen Unbekannten," *ibid.*, vol. 25 (1908), pp. 53-77.

[9] H. Weyl, "Ueber beschränkte quadratische Formen, deren Differenz vollstetig ist," *ibid.*, vol. 27 (1909), pp. 373-392.

[10] ———, "Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkürlicher Funktionen," *Mathematische Annalen*, vol. 68 (1910), pp. 222-269.

AN APPLICATION OF THE THEORY OF BASIC SERIES TO THEOREMS OF BERNSTEIN-WIDDER TYPE.*

By PHILIP DAVIS.

---

1. Introduction. That the signs of the derivatives of a function of class C^∞ on an interval may influence its analytic character was first pointed out by S. Bernstein [1]. Bernstein's theorem may be stated as follows.

THEOREM 1. *Let $f(x)$ be a real valued function defined and of class C^∞ in $-1 < x < 1$. If each derivative of $f(x)$ is non-negative in $-1 < x < 1$, then $f(x)$ is necessarily analytic in that interval.*

Some years later, a result of a similar nature, but this time showing the effect of the signs of the even derivatives, was obtained by D. V. Widder [2].

THEOREM 2. *If $f(x)$ is real and of class C^∞ in $-1 < x < 1$ and if*

$$(1) \qquad (-1)^n f^{(2n)}(x) \geqq 0 \qquad (-1 < x < 1;\ n = 0, 1, 2, \cdots),$$

then $f(x)$ necessarily coincides with an entire function of exponential type at most π.

Functions satisfying the condition (1) have been termed completely convex. Subsequently Boas and Pólya [3] gave a generalization which contained as special cases both Theorems 1 and 2. In the present note these theorems are generalized in another direction by obtaining similar results for a number of familiar sets of differential operators. We shall find it convenient to formulate these results using the notion of the basic series of polynomials introduced by Whittaker [4]. Our work is essentially algebraic in character, inasmuch as all questions are referred back either to Theorem 1 or to Theorem 2.

2. Basic sets and series. In Whittaker's terminology, a set of polynomials $\{p_n(z)\}$ is called a basic set if an arbitrary polynomial can be expressed in a unique fashion as a finite linear combination of the p_n's. Whittaker has given the following necessary and sufficient condition that a set of polynomials be basic.

---

* Received April 6, 1950.

THEOREM 3. *Given the set of polynomials*

$$p_n(z) = \sum_{j=0}^{\infty} p_{nj} z^j, \qquad (n = 0, 1, 2, \cdots) \tag{2}$$

where in (2) *we have inserted dummy zero coefficients; a necessary and sufficient condition that the set* (2) *be basic is that the row-finite matrix*

$$P = (p_{nj}) \qquad (n, j = 0, 1, 2, \cdots) \tag{3}$$

have an inverse Q *which is row-finite*:

$$PQ = QP = I. \tag{4}$$

By a row-matrix is meant one in which each row has but a finite number of non-zero elements. It may also be remarked here that the multiplication of row-finite matrices is associative, so that the above condition is equivalent to the existence of a unique Q for which $QP = I$.[1]

With a given basic set of polynomials (2) with matrix $P = (p_{nj})$ and with unique row-finite inverse $Q = (q_{nj})$, there may be associated the following set of (possibly formal) power series:

$$q_n(z) = \sum_{k=0}^{\infty} q_{kn} z^k / k! \qquad (n = 0, 1, 2, \cdots). \tag{5}$$

The associated set may be characterized in the following way. The set $\{q_n(z)\}$ is associated to the basic set $\{p_n(z)\}$ of polynomials if and only if

$$\sum_{n=0}^{\infty} p_n(z) q_n(w) = e^{zw} \tag{6}$$

holds as a formal identity, in the sense that when substitutions are made, order of summation reversed, and coefficients compared, we have equality. it is easily seen that as a formal identity, (6) is equivalent to (4).

We shall give a second characterization of the associated set. For two power series $p(z) = \sum_{n=0}^{\infty} a_n z^n$, $q(z) = \sum_{n=0}^{\infty} b_n z^n$, introduce the inner product

$$[p, q] = \sum_{n=0}^{\infty} a_n b_n n!. \tag{7}$$

If either p or q is a polynomial, there can be no question of the convergence of the series in (7). The set (5) is associated to the basic set (2) if and only if the two sets are biorthonormal:

$$[p_i, q_j] = \delta_{ij}. \tag{8}$$

---

[1] See, e. g., Dienes [5].

With a given function f (of an appropriate class) and a given basic set, there may be associated the following formal expansion of f:

$$f(z) \sim \sum_{n=0}^{\infty} [q_n, f] p_n(z), \tag{9}$$

the so-called basic series for f. The study of the representability of functions in a series (9) forms a major chapter in the theory of functions. For our purposes, however, we need only know that equality in (9) holds whenever f is a polynomial. Since $\{p_n(z)\}$ is basic, there is a unique finite expansion of a polynomial f: $f = c_0 p_0 + c_1 p_1 + \cdots + c_k p_k$. Hence by (7) and (8), $c_i = [q_i, f]$.

We shall now list a number of basic sets together with their associated sets. Our list is intended to be suggestive rather than exhaustive, and may be augmented at will. Moreover, the examples have been chosen so that the associated functions are expressible in closed form. Except when otherwise noted, the 0-th polynomial is 1.

Maclaurin Polynomials:

$$p_n(z) = z^n/n!; \qquad q_n(z) = z^n. \tag{10}$$

Newton Interpolation Polynomials:

$$p_n(z) = (z - a_0)(z - a_1) \cdots (z - a_{n-1}), \quad a_i \text{ distinct};$$

$$q_n(z) = \begin{vmatrix} 1 & 1 & \cdots 1 \\ a_0 & a_1 & \cdots a_n \\ & \cdots & \\ a_0^{n-1} & a_1^{n-1} \cdots & a_n^{n-1} \\ e^{a_0 z} & e^{a_1 z} & \cdots e^{a_n z} \end{vmatrix} : \begin{vmatrix} 1 & 1 & \cdots 1 \\ a_0 & a_1 & \cdots a_n \\ & \cdots & \\ a_0^{n} & a_1^{n} \cdots & a_n^{n} \end{vmatrix}. \tag{11}$$

Gontcharoff Polynomials (Whittaker [4]):

$$p_n(z) = \int_{a_0}^{z} dz_1 \int_{a_1}^{z_1} dz_2 \cdots \int_{a_{n-1}}^{z_{n-1}} dz_n, \qquad q_n(z) = z^n e^a n^z / n!. \tag{12}$$

Lidstone Polynomials (Widder [2]):

$$\begin{aligned} &p_0(z) = z, \qquad p_{2n+1}(z) = p_{2n}(1 - z); \\ &p_{2n-2}(z) = p_{2n}''(z), \qquad p_{2n}(0) = p_{2n}(1) = 0, \\ &q_{2n}(z) = z^{2n}, \qquad q_{2n+1}, z) = z^{2n} e^z. \end{aligned} \tag{13}$$

Hermite Polynomials:

$$(14)\quad p_n(z) = \sum_{k=0}^{[n/2]} (-1)^k (2z)^{n-2k}/k!(n-2k)!, \qquad q_n(z) = z^n e^{z^2/4}/2^n n!.$$

Appel Polynomials (Boas [6]):

$$A(w)e^{zw} = \sum_{n=0}^{\infty} p_n(z)w^n, \qquad q_n(z) = z^n/A(z),$$

$A(w)$ regular at $w = 0$, $A(0) \neq 0$.

3. Applications. We may now state our principal theorem.

THEOREM 4. *Let $\{p_n(z)\}$ be a basic set of polynomials with real coefficients, and let $\{q_n(z)\}$ be the (possibly formal) set of associated power series. For each n, let $[q_j(z), z^n] = q_{nj}$ have a constant sign as j varies. Then, with $f(x)$ a real valued function of class C^∞ on $-1 < x < 1$,*

$$(16)\quad p_n(D)f(x) \geqq 0 \qquad (n = 0, 1, 2, \cdots; -1 < x < 1; D \equiv d/dx)$$

implies that $f(x)$ is analytic in $-1 < x < 1$.

Proof. We have, for some integer k_n,

$$z^n = \sum_{j=0}^{k_n} [q_j(z), z^n]p_j(z); \text{ hence } D^n f = \sum_{j=0}^{k_n} q_{nj}p_j(D)f.$$

The inequality (16) now implies that $D^n f(x)$ has constant sign in $-1 < x < 1$. The theorem is therefore an immediate consequence of Theorem 1.

It may happen that the associated series $q_n(z)$ have non-negative coefficients for all n. In other words, the basic set $\{p_n(z)\}$ has a positive inverse property. All the sets of our list, with proper restrictions on the arbitrary constants appearing therein, have this property.

COROLLARY 4.1. *If $a_i \geqq 0$, then the sets* (10), (11), (12), (13), (14) *have a positive inverse property. This is also true of the set* (15) *provided that $[A(w)]^{-1}$ has non-negative Maclaurin coefficients. For the polynomials of any of these sets, condition* (16) *implies analyticity in $-1 < x < 1$.*

Proof. Inspection of the corresponding q_n's reveals that under the above conditions their coefficients are indeed non-negative. For the set (11) we should add that $q_{nj} = \Sigma a_0^{i_0} a_1^{i_1} \cdots a_j^{i_j}$ where the summation is extended over all non-negative i's for which $i_0 + i_1 + \cdots + i_j = n - j$.

Theorem 4 is capable of further generalization, for Theorem 1, as stated, is but a special case of Bernstein's results. Actually, Bernstein has shown

that if each successive difference of $f(x)$ is of constant sign in the interval, then, without assuming the existence of derivatives, it will follow that $f(x)$ is analytic in the interval. If therefore in Theorem 4, $p_n(D)$ be replaced by $p_n(\Delta)$, the conclusion follows without assuming the existence of derivatives.

By selecting the Newton Polynomials (11) and combining Theorems 4 and 2, we arrive at the following generalization of Widder's Theorem.

COROLLARY 4.2. *Let $f(x)$ be real and of class C^∞ on $-1 < x < 1$, and suppose that a_i are real. Then*

$$(17) \qquad f(x) \geqq 0, \qquad (-1)^n(D^2 + a_1^2) \cdots (D^2 + a_n^2) f(x) \geqq 0$$

$(n = 1, 2, \cdots; -1 < x < 1)$ implies that $f(x)$ is completely convex on the interval and hence coincides with an entire function of exponential type at most π.

HARVARD UNIVERSITY.

REFERENCES.

[1] S. Bernstein, "Sur la définition et les propriétés des fonctions analytiques d'une variable réelle," *Mathematische Annalen*, vol. 75 (1914), pp. 449-468.

[2] D. V. Widder, "Completely convex functions and Lidstone series," *Transactions of the American Mathematical Society*, vol. 51 (1942), pp. 387-398.

[3] R. P. Boas, Jr. and G. Pólya, "Influence of the signs of the derivatives of a function on its analytic character," *Duke Journal of Mathematics*, vol. 9 (1942), pp. 406-424.

[4] J. M. Whittaker, *Interpolatory Function Theory*, Cambridge Tracts No. 33, Cambridge, 1935.

[5] P. Dienes, "Notes on linear equations in infinite matrices," *Quarterly Journal of Mathematics*, vol. 3 (1932), pp. 253-268.

[6] R. P. Boas, Jr., "Exponential transforms and Appel polynomials," *Proceedings of the National Academy of Sciences*, vol. 34 (1948), pp. 481-483.

THE MEAN CONVERGENCE OF ORTHOGONAL SERIES.*

By G. Milton Wing.

1. Introduction. We consider a sequence of functions $\{\phi_n(x)\}$, $n = 0, 1, 2, \cdots$, orthonormal with respect to a measure

$$\int_a^b \phi_n(x)\phi_m(x)\,dF(x) = \delta_{nm},$$

and the corresponding formal expansion of an arbitrary function in terms of this sequence,

$$f(x) \sim \sum_{n=0}^{\infty} a_n\phi_n(x), \tag{1.1}$$

where $a_n = \int_a^b f(x)\phi_n(x)\,dF(x)$. While the local convergence properties of series of the type (1.1) have been thoroughly investigated for particular sets of $\phi_n(x)$, their mean convergence has been studied but little. This problem takes the form: for what values of p, $1 \leqq p < \infty$ does the existence of the integral $\int_a^b |f(x)|^p\,dF(x)$ imply that

$$\lim_{N\to\infty} \int_a^b \left| f(x) - \sum_{n=0}^{N} a_n\phi_n(x) \right|^p dF(x) = 0? \tag{1.2}$$

When equation (1.2) holds for every $f(x)$ such that $\int_a^b |f(x)|^p\,dF(x)$ exists we shall say that the sequence $\{\phi_n(x)\}$ forms a *basis* for the space of these functions.

In the case that $F(x)$ is an increasing, absolutely continuous function a different question may be posed. If $dF(x) = w(x)\,dx$ then the sequence $\{\phi_n(x)w^{\frac{1}{2}}(x)\}$ is orthonormal in the classical sense and one is led to the formal expansion

$$f(x) \sim \sum_{n=0}^{\infty} b_n\phi_n(x)w^{\frac{1}{2}}(x), \tag{1.3}$$

* Received November 16, 1949; revised April 2, 1950. Presented to the Society October 29, 1949. The research was carried on while the author held an Atomic Energy Commission Predoctoral Fellowship.

where $b_n = \int_a^b f(x)\phi_n(x)w^{\frac{1}{2}}(x)\,dx$. We now ask: for what values of p, $1 \leqq p < \infty$, does the measurability of $f(x)$ and the relation $\int_a^b |f(x)|^p\,dx < \infty$ (i. e., $f(x)$ in L^p) imply

$$\lim_{N\to\infty} \int_a^b \Big| f(x) - \sum_{n=0}^{N} b_n\phi_n(x)w^{\frac{1}{2}}(x)\Big|^p\,dx = 0?$$

Problems of this sort were first investigated by M. Riesz [15], who considered the case of trigonometric functions, and later by Schauder, Kober, Caton and Hille, and Paley [16, 8, 4, 10], who investigated other special sets of functions. Recently, H. Pollard has studied extensively the mean convergence of orthogonal polynomials [11, 12, 13, 14]. We shall make some additions to his results and apply similar methods to the problems associated with expansions in series of Bessel functions. It should be noted that the L^p theory of Hankel transforms (the integral analogue of Fourier-Bessel Series) has been studied by Busbridge [3].

2. General principles. Our interest will be confined to a sequence $\{\phi_n(x)\}$ orthonormal over a finite interval with respect to a non-negative weight function $w(x)$. We shall assume the elements of the sequence belong to L_w^p (i. e. $\int_a^b |\phi_n(x)|^p w(x)\,dx < \infty$ for all n) and to the conjugate space $L_w^{p'}$. We assume also that the sequence is closed over both of these spaces. The notation $\|f(x)\|_p$ will regularly be used to mean $\left\{\int_a^b |f(x)|^p\,dx\right\}^{1/p}$ and $S_N(f,x)$ will denote the N-th partial sum of the series in (1.1). The following principles will be used repeatedly.

(i) To establish that $\{\phi_n(x)\}$ is a basis for L_w^p, $p > 1$, it suffices to prove that, for all $f(x)$ in L_w^p,

$$\overline{\lim_{N\to\infty}} \| S_N(f,x)w^{1/p}(x)\|_p < +\infty.$$

The proof of this has been given by Pollard [13, p. 361].

(ii) If $\{\phi_n(x)\}$ is a basis for L_w^p, $p > 1$, then it is also a basis for $L_w^{p'}$.

This principle may be found in Banach [1, p. 108] for the case $w(x) \equiv 1$. His proof may be generalized to include our case.

CHAPTER I. BESSEL SERIES.

3. Orthogonality of sets of Bessel functions. It is well known [19, p. 134] that if $\{\lambda_n\}$ $(n = 1, 2, \cdots)$ is the sequence of successive positive roots of the equation

$$xJ_\nu'(x) + HJ_\nu(x) = 0, \tag{3.1}$$

where H is a real number and $\nu \geqq -1$, then

$$\int_0^1 xJ_\nu(\lambda_n x)J_\nu(\lambda_m x)\,dx = \tfrac{1}{2}\{(1 - \nu^2/\lambda^2_n)J_\nu^2(\lambda_n) + J_\nu'^2(\lambda_n)\}\delta_{nm}. \tag{3.2}$$

In case $H = +\infty$ the sequence, which we now denote by $\{\mu_n\}$, becomes the sequence of positive roots of $J_\nu(x)$. In place of (3.2) we have

$$\int_0^1 xJ_\nu(\mu_n x)J_\nu(\mu_m x)\,dx = \tfrac{1}{2}J_{\nu+1}^2(\mu_n)\delta_{nm}. \tag{3.3}$$

Series based upon the relation (3.3) are usually referred to as Fourier-Bessel series. They are somewhat easier to handle than the more general Dini series arising from equation (3.2). Fortunately, however, the convergence properties of the latter may be deduced from those of the simpler series.

4. Fourier-Bessel series of the type (1.3). We consider the sequence $\{(2x)^{\frac{1}{2}}J_\nu(\mu_n x)/J_{\nu+1}(\mu_n)\}$, and write

$$S_N(f, x) = \sum_{n=1}^{N} b_n(2x)^{\frac{1}{2}}J_\nu(\mu_n x)/J_{\nu+1}(\mu_n),$$

where $J_{\nu+1}(\mu_n)b_n = 2^{\frac{1}{2}}\int_0^1 t^{\frac{1}{2}}J_\nu(\mu_n t)f(t)\,dt.$

Then we have $S_N(f, x) = \int_0^1 (xt)^{\frac{1}{2}}U_N(t, x)f(t)\,dt$, where

$$U_N(t, x) = 2\sum_{n=1}^{N} J_\nu(\mu_n x)J_\nu(\mu_n t)/J_{\nu+1}^2(\mu_n).$$

The following lemma, due to MacRobert [9], is crucial in the proofs of many of the theorems of this chapter.

LEMMA 4.1. *If* $\nu \geqq -\frac{1}{2}$ *and* $A_N = (N + \frac{1}{2}\nu + \frac{1}{4})\pi$, *then*

$$|U_N(t, x) - \tfrac{1}{2}A_N\{J_\nu(A_N x)J_{\nu+1}(A_N t) - J_\nu(A_N t)J_{\nu+1}(A_N x)\}/(x - t)|$$
$$< C(xt)^{-\frac{1}{4}}\{(x + t)^{-1} + (2 - x - t)^{-1}\},$$

for all x *and all* t *in* $[0, 1]$.

This result we now use to obtain our first theorem concerning mean convergence of Bessel series.

THEOREM 4.1. *If* $\nu \geqq -\frac{1}{2}$ *and* $f(x)$ *is in* L^p *for* $p > 1$ *then*

$$\lim_{N\to\infty} \int_0^1 |f(x) - S_N(f,x)|^p \, dx = 0.$$

Remark. We note the relations $J_{\frac{1}{2}}(x) = (2/\pi x)^{\frac{1}{2}} \sin x$ and $J_{-\frac{1}{2}}(x) = (2/\pi x)^{\frac{1}{2}} \cos x$. Hence when $\nu = \pm \frac{1}{2}$ our series become trigonometric series, and our theorem therefore includes the classical result of M. Riesz on mean convergence of these series [15].

Proof of the theorem. Since the function $t^{-\nu}J_\nu(t)$ is entire [see 19, p. 40], the condition $\nu \geqq -\frac{1}{2}$ implies that the functions $t^{\frac{1}{2}}J_\nu(\mu_n t)$ are bounded in t in $[0,1]$. The coefficients a_n are therefore well-defined. The lemma suggests that we write

$$\begin{aligned} S_N(f,x) = \int_0^1 \{ & U_N(t,x) \\ & - \tfrac{1}{2}A_N[J_\nu(A_Nx)J_{\nu+1}(A_Nt) - J_\nu(A_Nt)J_{\nu+1}(A_Nx)]/(x-t) \\ & + \tfrac{1}{2}A_N[J_\nu(A_Nx)J_{\nu+1}(A_Nt) - J_\nu(A_Nt)J_{\nu+1}(A_Nx)]/(x-t)\}(xt)^{\frac{1}{2}}f(t)\,dt. \end{aligned}$$

From the lemma we deduce that

$$\begin{aligned} (4.1) \quad |S_N(f,x)| \leqq C \int_0^1 & \{(x+t)^{-1} + (2-x-t)^{-1}\}|f(t)|\,dt \\ & + \tfrac{1}{2}\left| \int_0^1 (A_Nx)^{\frac{1}{2}}J_\nu(A_Nx)(A_Nt)^{\frac{1}{2}}J_{\nu+1}(A_Nt)f(t)/(x-t)\,dt \right| \\ & + \tfrac{1}{2}\left| \int_0^1 (A_Nx)^{\frac{1}{2}}J_{\nu+1}(A_Nx)(A_Nt)^{\frac{1}{2}}J_\nu(A_Nt)f(t)/(x-t)\,dt \right|. \end{aligned}$$

The convergence almost everywhere of the improper integrals, as Cauchy principal values, follows from the fact that the functions $(A_Nt)^{\frac{1}{2}}J_\nu(A_Nt)f(t)$ and $(A_Nt)^{\frac{1}{2}}J_{\nu+1}(A_Nt)f(t)$ are integrable [18, p. 132].

We shall now prove that

$$(4.2) \qquad \overline{\lim_{N\to\infty}} \| S_N(f,x) \|_p < +\infty.$$

By Hilbert's inequality [6, p. 226] the function $g(x) = \int_0^1 |f(t)|/(x+t)dt$

is in L^p so that $\| g(x) \|_p < \infty$. A simple change of variables allows us to make a similar statement concerning $\int_0^1 | f(t) | / (2 - x - t)\, dt$.

The remaining two integrals of (4. 1) are similar, and we need examine only one of them. We follow the method used by M. Riesz in his treatment of trigonometric series [15]. Write

$$I_N(x) = \int_0^1 (A_N x)^{\frac{1}{2}} J_\nu(A_N x)(A_N t)^{\frac{1}{2}} J_{\nu+1}(A_N t) f(t)/(x - t)\, dt$$
$$= B_N(x) \int_0^1 C_N(t) f(t)/(x - t)\, dt,$$

where $B_N(x) = (A_N x)^{\frac{1}{2}} J_\nu(A_N x)$, $C_N(t) = (A_N t)^{\frac{1}{2}} J_{\nu+1}(A_N t)$.

From the asymptotic form of the Bessel functions [19, p. 194], we conclude that $B_N(x)$ and $C_N(t)$ are bounded in N, x, and t. Hence $C_N(t) f(t)$ is a member of L^p. It follows by a theorem of Riesz [15] that $I_N(x)$ is also in L^p and $\| I_N(x) \|_p < A \| C_N(x) f(x) \|_p < A \| f(x) \|_p < A$, since B_N and C_N are bounded.

We now apply Minkowski's inequality to (4. 1) to get

$$\| S_N(f, x) \|_p \leqq C \left\| \int_0^1 \{(x + t)^{-1} + (2 - x - t)^{-1}\} | f(t) |\, dt \right\|_p$$
$$+ \tfrac{1}{2} \left\| B_N(x) \int_0^1 C_N(t) f(t)/(x - t)\, dt \right\|_p$$
$$+ \tfrac{1}{2} \left\| C_N(x) \int_0^1 B_N(t) f(t)/(x - t)\, dt \right\|_p < +\infty.$$

Hence (4. 2) is established.

If it can now be shown that the sequence $\{(2x)^{\frac{1}{2}} J_\nu(\mu_n x)/J_{\nu+1}(\mu_n)\}$ is closed in L^p, principle (i) of § 2 may be applied to complete the proof. W. H. Young has proved if $t^{\frac{1}{2}} f(t)$ is in L, then the vanishing of $\int_0^1 t f(t) J_\nu(\mu_n t)\, dt$ for all n implies that $f(t) \equiv 0$ [21, pp. 174-175]. From this follows the completeness of our sequence in L^p, and hence its closure [see 7, p. 200].

5. Fourier-Bessel series of the type (1. 1). We turn next to series of type (1. 1) with $\phi_n(x) = 2^{\frac{1}{2}} J_\nu(\mu_n x)/J_{\nu+1}(\mu_n)$ and

$$J_{\nu+1}(\mu_n) a_n = 2^{\frac{1}{2}} \int_0^1 t f(t) J_\nu(\mu_n t)\, dt.$$

The partial sums have the form $T_N(f,x) = \int_0^1 tf(t)U_N(t,x)\,dt$. We prove the following theorem.

THEOREM 5.1. *If $\nu \geqq -\frac{1}{2}$ and $f(t)$ is in L_t^p, $4/3 < p < 4$, then*

$$\lim_{N\to\infty} \int_0^1 |f(x) - T_N(f,x)|^p x\,dx = 0.$$

Proof. The coefficients a_n exist provided $t^{\frac{1}{2}}|f(t)|$ is in L, since $t^{\frac{1}{2}}J_\nu(\mu_n t)$ is bounded in t in $[0,1]$. Now

$$\int_0^1 t^{\frac{1}{2}}|f(t)|\,dt \leqq \left\{\int_0^1 t^{(\frac{1}{2}-1/p)p'}dt\right\}^{1/p'} \left\{\int_0^1 t\,|f(t)|^p dt\right\}^{1/p}.$$

The last integral exists by assumption. The first on the right converges provided $(\frac{1}{2} - 1/p)p' = (1/p' - \frac{1}{2})p' > -1$ or $p' < 4$, as assumed.

Proceeding exactly as in Theorem 4.1, we find that

$$(5.1)\quad |T_N(f,x)| \leqq Cx^{-\frac{1}{2}}\int_0^1 t^{\frac{1}{2}}|f(t)|\{(x+t)^{-1} + (2-x-t)^{-1}\}dt$$
$$+\tfrac{1}{2}\left|\int_0^1 A_N J_\nu(A_N x)J_{\nu+1}(A_N t)tf(t)/(x-t)\,dt\right|$$
$$+\tfrac{1}{2}\left|\int_0^1 A_N J_\nu(A_N t)J_{\nu+1}(A_N x)tf(t)/(x-t)\,dt\right|.$$

This expression must now be treated more carefully than the similar one for $|S_N(f,x)|$ encountered in § 4. Our goal is to prove

$$(5.2)\qquad \overline{\lim_{N\to\infty}} \|T_N(f,x)x^{1/p}\|_p < +\infty.$$

We begin by considering

$$(5.3)\quad g(x) = x^{1/p-\frac{1}{2}}\int_0^1 t^{\frac{1}{2}}|f(t)|/(x+t)\,dt$$
$$= \int_0^1 t^{1/p}|f(t)|/(x+t)dt + \int_0^1 t^{1/p}|f(t)|\{(t/x)^{\frac{1}{2}-1/p} - 1\}/(x+t)dt$$

$= I_1(x) + I_2(x)$. Hilbert's inequality again applies to $I_1(x)$. By the converse of Hölder's inequality [6, p. 120], $I_2(x)$ belongs to L^p if the double integral

$$(5.4)\qquad I = \int_0^1 h(x)\,dx \int_0^1 t^{1/p}\,|f(t)|\{(t/x)^{\frac{1}{2}-1/p}-1\}/(x+t)\,dt$$

is bounded for every $h(x)$ in $L^{p'}$.

Applying Hölder's inequality for double integrals to (5.4) we have

$$|I| = \left|\int_0^1 dx \int_0^1 h(x)\{[(t/x)^{\frac{1}{2}-1/p}-1]/(x+t)\}^{1/p'} t^{1/p}\,|f(t)|\,(t/x)^{-1/pp'}\right.$$
$$\left.\cdot\{[t/x)^{\frac{1}{2}-1/p}-1]/(x+t)\}^{1/p}(t/x)^{1/pp'}\,dt\right|$$
$$\leqq \left\{\int_0^1 |h(x)|^{p'}dx \int_0^1 |(t/x)^{\frac{1}{2}-1/p}-1\,|\,(t/x)^{-1/p}/(x+t)\,dt\right\}^{1/p'}$$
$$\cdot\left\{\int_0^1 t\,|f(t)|^p dt \int_0^1 |(t/x)^{\frac{1}{2}-1/p}-1\,|\,(t/x)^{1/p'}/(x+t)\,dx\right\}^{1/p}.$$

In view of the conditions imposed on p, it now suffices to prove the boundedness, for all x in $(0,1)$, of

$$J_1 = \int_0^1 |(t/x)^a - (t/x)^b\,|/(x+t)\,dt,$$

where $0 > a > -1$, $0 > b > -1$. The trivial substitution $z = tx$ yields

$$J_1 = \int_0^{1/x} |z^a - z^b\,|/(1+z)\,dz \leqq \int_0^\infty |z^a - z^b\,|/(1+z)\,dz.$$

The last integral is clearly finite. Hence I converges and $g(x)$ is in L^p. Similarly, $g_1(x) = x^{1/p-\frac{1}{2}}\int_0^1 t^{\frac{1}{2}}\,|f(t)|/(2-x-t)\,dt$ is a member of L^p.

Finally we turn to the last two integrals of (5.1). It suffices to consider just one of these. Let

$$G_N(x) = x^{1/p}\int_0^1 A_N J_\nu(A_N x) J_{\nu+1}(A_N t)\,tf(t)/(x-t)\,dt$$
$$= x^{1/p-\frac{1}{2}} B_N(x)\int_0^1 C_N(t)\,t^{\frac{1}{2}} f(t)/(x-t)\,dt.$$

In proving that $G_N(x)$ is in L^p, it will do no harm to suppose $f(t) \geqq 0$. Write

$$G_N(x) = B_N(x)\int_0^1 C_N(t)\,t^{1/p} f(t)/(x-t)\,dt$$
$$+ B_N(x)\int_0^1 C_N(t)\,t^{1/p} f(t)\{(t/x)^{\frac{1}{2}-1/p}-1\}/(x-t)\,dt.$$

The first integral is covered by Riesz's theorem. The second may be treated by the device used in investigating $I_2(x)$ of (5.3) and leads to

$$J_2 = \int_0^1 |(t/x)^a - (t/x)^b|/(x-t)\,dt.$$

This is finite for x in $(0,1)$. Hence

$$\|G_N(x)\|_p \leqq A \left\| \int_0^1 C_N(t) t^{1/p} f(t)/(x-t)\,dt \right\|_p + A \left\| \int_0^1 C_N(t) t^{1/p} f(t)\{(t/x)^{\frac{1}{2}-1/p} - 1\}/(x-t)\,dt \right\|_p \leqq A.$$

Here A is independent of N since the B's are bounded in N. This establishes (5.2).

We need now only establish the closure of the sequence $\{J_\nu(\mu_n t)\}$ in L_t^p. Since $f(t)$ is in L_t^p for $p > 4/3$ it follows that $t^{\frac{1}{2}} f(t)$ is in L. Hence we may again use the result of Young and assert that the sequence is complete in L_t^p for $p > 4/3$, and so closed there for $p < 4$.

6. A counterexample. We have seen that the condition $p > 4/3$ is necessary to insure the existence of the coefficients a_n, provided we allow the order ν of the Bessel functions to be as small as $-\frac{1}{2}$. Our argument depended upon the fact that $t^{\frac{1}{2}} J_\nu(\mu_n t)$ is bounded in t, $0 \leqq t \leqq 1$, for $\nu \geqq -\frac{1}{2}$. Since, more generally, $t^{-\nu} J_\nu(\mu_n t)$ is bounded in t, one might expect that the conditions on p should really depend upon ν. The following example shows that this is not the case. In view of principle (ii) of § 2, we may confine our attention to the interval $1 \leqq p < 4/3$.

THEOREM 6.1. *There exists a function belonging to L_t^p, $1 \leqq p < 4/3$, whose Fourier-Bessel series does not converge in L_t^p.*

Proof. The function $f(t) = t^{-3/2}$ belongs to L_t^p for $p < 4/3$. To prove that its Fourier-Bessel series does not converge in the mean it suffices to prove that

$$\overline{\lim_{n\to\infty}} |a_n| \left\{ \int_0^1 t\,|J_\nu(\mu_n t)|^p\,dt \right\}^{1/p} > 0,$$

where

$$J_{\nu+1}^2(\mu_n) a_n = 2 \int_0^1 t^{-\frac{1}{2}} J_\nu(\mu_n t)\,dt. \tag{6.1}$$

We may confine our attention to the case $\nu > -\frac{1}{2}$ since the coefficients a_n

do not even exist when $\nu = -\frac{1}{2}$. Substituting z for $t\mu_n$ in (6.1) we have [see 20, p. 383]

$$(6.2)\quad \lim_{n\to\infty} a_n\mu_n^{\frac{1}{2}}J_{\nu+1}^2(\mu_n) = 2\int_0^\infty z^{-\frac{1}{2}}J_\nu(z)\,dz = 2^{-\frac{1}{2}}\Gamma(\tfrac{1}{2}\nu+\tfrac{1}{4})/\Gamma(\tfrac{1}{2}\nu+\tfrac{3}{4})$$

$= A > 0$. Now

$$(6.3)\qquad \int_0^1 t\,|\,J_\nu(\mu_n t)\,|^p\,dt = \mu_n^{-2}\int_0^{\mu_n} z\,|\,J_\nu(z)\,|^p\,dz > B_p\mu_n^{-p/2},$$

as one may verify by use of the asymptotic formula

$$J_\nu(z) \sim (2/\pi z)^{\frac{1}{2}}\cos(z - \tfrac{1}{2}\nu\pi - \tfrac{1}{4}\pi)$$

[19, p. 195]. Combining (6.2) and (6.3) we get

$$\overline{\lim_{n\to\infty}}\,|\,a_n\,|\,\left\{\int_0^1 t\,|\,J_\nu(\mu_n t)\,|^p\,dt\right\}^{1/p} > \overline{\lim_{n\to\infty}}\,2AB_p/\mu_n J_{\nu+1}^2(\mu_n) > 0,$$

as desired.

It is interesting to note that the interval of mean convergence, $4/3 < p < 4$, is the same as that for Legendre series and for series of more general types of polynomials to be discussed in § **9**. In all of these cases the question of convergence at the endpoints of the interval is still open.

7. Dini series. From (3.2) it is clear that the functions $k_n x^{\frac{1}{2}}J_\nu(\lambda_n x)$ where $2/k_n^2 = (1-\nu^2/\lambda_n^2)J_\nu^2(\lambda_n) + J_\nu'^2(\lambda_n)$ form an orthonormal sequence over $[0,1]$ when $\nu \geqq -1$. We are led to consider the formal expansion

$$(7.1)\qquad f(x) \sim \sum_{n=0}^{\infty} k_n\beta_n x^{\frac{1}{2}}J_\nu(\lambda_n x),$$

where $\beta_n = k_n\int_0^1 f(t)t^{\frac{1}{2}}J_\nu(\lambda_n t)\,dt$. Dini has shown in his study of pointwise convergence of such series that (7.1) is not always the correct formal expansion [5]. Rather, one must add to the series of (7.1) a term $B_0(x)$ where

$B_0(x) = 0$ if $H+\nu > 0$; $B_0(x) = 2(\nu+1)x^{\nu+\frac{1}{2}}\int_0^1 t^{\nu+\frac{1}{2}}f(t)\,dt$ if $H+\nu = 0$;

$B_0(x) = 2\lambda_0^2 I_\nu(\lambda_0 x)x^{\frac{1}{2}}/\{(\lambda_0^2+\nu^2)I_\nu^2(\lambda_0) - \lambda_0^2 I_\nu'^2(\lambda_0)\}\int_0^1 t^{\frac{1}{2}}f(t)I_\nu(\lambda_0 t)\,dt$ if $H+\nu < 0$.

Here $\pm i\lambda_0$ are the two purely imaginary zeros of (3.1) known to exist when

$H+\nu<0$ and $I_\nu(x)=e^{-\frac{1}{2}\nu\pi i}J_\nu(xe^{\frac{1}{2}\pi i})$. Denoting the partial sums of the amended series by $\sigma_N(f,x)$, we have $\sigma_N(f,x)=\int_0^1 (xt)^{\frac{1}{2}}C_N(t,x)f(t)\,dt$, where

$$C_N(t,x)=A_0(x,t)+\sum_{n=1}^{N}k_n^2J_\nu(\lambda_n x)J_\nu(\lambda_n t),$$

$A_0(x,t)=0$ if $H+\nu>0$; $A_0(x,t)=2(\nu+1)x^\nu t^\nu$ if $H+\nu=0$;

$$A_0(x,t)=2\lambda_0^2 I_\nu(\lambda_0 x)I_\nu(\lambda_0 t)/\{(\lambda_0^2+\nu^2)I_\nu^2(\lambda_0)-\lambda_0^2 I_\nu'^2(\lambda_0)\}$$

if $H+\nu<0$.

We now prove the following theorem.

THEOREM 7.1. *If $\nu\geqq-\frac{1}{2}$ and $f(x)$ is in L^p, $p>1$, then*

$$\lim_{N\to\infty}\int_0^1 |f(x)-\sigma_N(f,x)|^p\,dx=0.$$

Proof. Using the notation of the previous section we let μ_m be the largest zero of $J_\nu(x)$ which is less than λ_{N+1}. Then it may be shown that

$$|C_N(t,x)-U_m(t,x)|<C/(xt)^{\frac{1}{2}}(2-x-t)$$ for all x and t in $[0,1]$

[19, p. 599]. Hence

$$|\sigma_N(f,x)|=\left|\int_0^1 (xt)^{\frac{1}{2}}f(t)\{C_N(t,x)-U_m(t,x)+U_m(t,x)\}dt\right|$$

$$\leqq C\int_0^1 |f(t)|/(2-x-t)\,dt+|S_m(f,x)|.$$

The arguments of § 4 yield

$$\|\sigma_N(f,x)\|_p<A_p+\|S_m(f,x)\|_p<A_p.$$

To complete the proof we must establish the closure in L^p, $p>1$, of the orthogonal sequence $\{\phi_n(x)\}$, where $\phi_n(x)=x^{\frac{1}{2}}J_\nu(\lambda_n x)$ for $n>0$, and $\phi_0(x)=0$, $x^{\nu+\frac{1}{2}}$ or $x^{\frac{1}{2}}I_\nu(\lambda_0 x)$, according as $H+\nu>0$, $H+\nu=0$, or $H+\nu<0$. Young has proved the completeness of $\{\phi_n(x)\}$ and the closure then follows. This establishes the theorem.

To complete this section we write $\alpha_n=k_n\int_0^1 tf(t)J_\nu(\lambda_n t)\,dt$, and consider $\tau_N(f,x)=C_0(x)+\sum_{n=1}^{N}\alpha_n k_n J_\nu(\lambda_n x)$, where

$C_0(x) = 0$ if $H + \nu > 0$; $C_0(x) = 2(\nu+1)x^\nu \int_0^1 t^{\nu+1} f(t)\,dt$ if $H + \nu = 0$;

$C_0(x) = 2\lambda_0^2 I_\nu(\lambda_0 x)/\{(\lambda_0^2 + \nu^2) I_\nu^2(\lambda_0) - \lambda_0^2 I_\nu'^2(\lambda_0)\} \int_0^1 t f(t) I_\nu(\lambda_0 t)\,dt$

if $H + \nu < 0$.

The methods we have employed in this paper now suffice to prove the following theorem.

THEOREM 7.2. *If $\nu \geqq -\frac{1}{2}$ and $f(t)$ is in L_t^p, $4/3 < p < 4$, then*

$$\lim_{N\to\infty} \int_0^1 |f(x) - \tau_N(f, x)|^p x\,dx = 0. \tag{7.2}$$

Moreover, there exists a function in L_t^p, $1 \leqq p < 4/3$ (or $p > 4$) such that (7.2) fails to hold.

CHAPTER II. ORTHOGONAL POLYNOMIALS.

8. In this chapter we shall extend some of the results of H. Pollard on the problem of mean convergence of series or orthogonal polynomials. The first theorem, without proof, has been announced in a special case by Pollard [11]. The second improves one of his results, while the last make it possible to absorb a wide class of polynomials into the existing theory.

9. Extension of a theorem of Pollard. The following theorem is due to Pollard in the special case when $w(x) = (1-x)^\alpha(1+x)^\beta$, $\alpha \geqq -\frac{1}{2}$, $\beta \geqq -\frac{1}{2}$.

THEOREM 9.2. *Let the polynomials $\{p_n(x)\}$ be orthonormal on $[-1, 1]$ with weight function $w(x)$. Let $(1-x^2)^{-1}q_n(x)$ denote the polynomials orthonormal with respect to the weight function $w(x)(1-x^2)$ and let the following conditions hold* [see 13, p. 356]:

(H-1) $w(x) \geqq 0$ *almost everywhere on* $[-1, 1]$;

(H-2) $\int_{-1}^1 |\log w(x)|\,(1-x^2)^{-\frac{1}{2}}\,dx < \infty$;

(H-3) $(1-x^2)^{\frac{1}{4}}\,|p_n(x)|\,w^{\frac{1}{2}}(x) \leqq A$ *for all x in $[-1, 1]$ and all $n = 0, 1, 2, \cdots$;*

(H-4) $(1-x^2)^{-\frac{1}{4}}\,|q_n(x)|\,w^{\frac{1}{2}}(x) \leqq A$ *for all x in $[-1, 1]$ and all $n = 0, 1, 2, \cdots$.*

Then, if $f(t)$ *is in* L^p, $4/3 < p < 4$, *and* $b_n = \int_{-1}^{1} f(x)p_n(x)w^{\frac{1}{2}}(x)\,dx$, *we have*

$$\lim_{N\to\infty} \int_{-1}^{1} |f(x) - \sum_{n=0}^{N} b_n w^{\frac{1}{2}}(x)p_n(x)|^p\,dx = 0. \tag{9.1}$$

Proof. We first show that the coefficients b_n are uniformly bounded. For, using (H-3),

$$|b_n| = \left|\int_{-1}^{1} f(x)p_n(x)w^{\frac{1}{2}}(x)\,dx\right| \leqq \left\{\int_{-1}^{1} |f(x)|^p\,dx\right\}^{1/p} \cdot \left\{\int_{-1}^{1} |p_n(x)w^{\frac{1}{2}}(x)|^{p'}\,dx\right\}^{1/p'}$$

$$\leqq A\,\|f(x)\|_p \left\{\int_{-1}^{1} (1-x^2)^{-p'/4}dx\right\}^{1/p'} < A_p, \text{ since } p' < 4.$$

We now write $S_N(f,x) = \sum_{n=0}^{N} b_n w^{\frac{1}{2}}(x)p_n(x)$, and seek to prove that $\overline{\lim_{N\to\infty}} \|S_N(f,x)\|_p < \infty$. Substituting the value of b_n and rearranging somewhat, we have

$$S_N(f,x) = w^{\frac{1}{2}}(x) \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)k_N(x,t)\,dt.$$

Here [see 13, p. 358].

$$k_N(x,t) = \alpha_N p_{N+1}(x)\,q_N(t)/(x-t) + \beta_N q_N(x)p_{N+1}(t)/(x-t) + \gamma_N p_{N+1}(x)p_{N+1}(t),$$

the constants α_N, β_N, γ_N being bounded in N. Therefore

$$S_N(f,x) = \alpha_N p_{N+1}(x)w^{\frac{1}{2}}(x) \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)q_N(t)/(x-t)\,dt + \beta_N q_N(x)w^{\frac{1}{2}}(x) \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)p_{N+1}(t)/(x-t)\,dt + \gamma_N p_{N+1}(x)w^{\frac{1}{2}}(x) \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)p_{N+1}(t)\,dt.$$

The last integral is just b_{N+1} and so is bounded. (H-3) and (H-4) then yield

$$\|S_N(f,x)\|_p < A \left\| (1-x^2)^{-\frac{1}{4}} \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)q_N(t)/(x-t)\,dt \right\|_p + A \left\| (1-x^2)^{\frac{1}{4}} \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)p_{N+1}(t)/(x-t)\,dt \right\|_p + A. \tag{9.2}$$

We now write $q_N(t)w^{\frac{1}{2}}(t) = A_N(t)(1-t^2)^{\frac{1}{4}}$, where $A_N(t)$ is bounded in N and t. Hence

$$\left\| (1-x^2)^{-\frac{1}{4}} \int_{-1}^{1} f(t)w^{\frac{1}{2}}(t)q_N(t)/(x-t)\,dt \right\|_p$$

$$\leqq \left\| \int_{-1}^{1} A_N(t)f(t)/(x-t)\,dt \right\|_p$$

$$+ \left\| \int_{-1}^{1} A_N(t)f(t)\{(1-t^2)^{\frac{1}{4}}(1-x^2)^{-\frac{1}{4}}-1\}/(x-t)\,dt \right\|_p.$$

According to the theorem of M. Riesz [15] the first term is bounded. To settle the second we need only prove that

$$\int_{-1}^{1} |f(t)| \left| (1-t^2)^{\frac{1}{4}}(1-x^2)^{-\frac{1}{4}} - 1 \right| / |x-t|\,dt$$

is in L^p. This follows at once from a result of Pollard [13, p. 363]. The second integral of (9.2) is handled similarly.

The proof is now completed by noting that the sequence $\{w^{\frac{1}{2}}(x)p_n(x)\}$ is closed in L^p.

10. Weight functions of class B. We shall say that the weight function $w(x) = t(x)(1-x)^{\alpha}(1+x)^{\beta}$ $(\alpha \geqq -\frac{1}{2}, \beta \geqq -\frac{1}{2})$ belongs to the class **B** if $t(x)$ is positive on $[-1,1]$ and has there a continuous derivative with the property that $t'(x+h) - t'(x) = O(n^{-2}|h|)$ as $h \to 0$ uniformly in $[-1,1]$.

The class **B** has been studied extensively by S. Bernstein. It may be shown that conditions (H-1) thru (H-4) of Theorem 9.1 are satisfied for a sequence orthonormal with respect to a member of **B** [2, pp. 267-268; 13, p. 362]. The case $t(x) \equiv 1$ yields the Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ [17, p. 57]. We now prove the following theorem concerning mean convergence of Jacobi series.

THEOREM 10.1. *Let* $w(x) = (1-x)^{\alpha}(1+x)^{\beta}$ $(\alpha \geqq -\frac{1}{2}, \beta \geqq -\frac{1}{2})$. *Then, if* $f(x)$ *is in* L^p, $4/3 < p < 4$, *equation* (9.1) *holds. Further, for* $1 \leqq p < 4/3$ *(or* $p > 4$*), there exists a function in* L^p *for which* (9.1) *fails to hold.*

Proof. The first part of the theorem follows at once from Theorem 9.1. For the second part it suffices to find a function $f(x)$ in L^p, $1 \leqq p < 4/3$, such that

$$\overline{\lim_{n\to\infty}} |\, b_n \,| \left\{ \int_{-1}^{1} |(1-x)^{\frac{1}{2}\alpha}(1+x)^{\frac{1}{2}\beta} P_n^{(\alpha,\beta)}(x)|^p \, dx \right\}^{1/p} > 0.$$

Here [see 17, p. 238]

$$b_n h_n^{(\alpha,\beta)} = \int_{-1}^{1} f(x)\,(1-x)^{\frac{1}{2}\alpha}(1+x)^{\frac{1}{2}\beta} P_n^{(\alpha,\beta)}(x)\,dx.$$

The function $f(x) = (1+x)^{\frac{1}{2}\beta}(1-x)^{-3/4}$ is in L^p and [see 17, p. 249],

$$b_n h_n^{(\alpha,\beta)} = \int_{-1}^{1} (1+x)^{\beta}(1-x)^{\frac{1}{2}\alpha - 3/4} P_n^{(\alpha,\beta)}(x)\,dx \sim n^{\frac{1}{2}}. \tag{10.1}$$

Further, $I = \int_0^1 (1-x)^{\frac{1}{2}\alpha} \,|\, P_n^{(\alpha,\beta)}(x) \,|\, dx$

$$= \int_0^1 (1-x)^{\frac{1}{2}\alpha}(1+x)^{\frac{1}{2}\beta}(1+x)^{-\frac{1}{2}\beta} \,|\, P_n^{(\alpha,\beta)}(x) \,|\, dx$$

$$\leqq \left\{ \int_{-1}^{1} |(1-x)^{\frac{1}{2}\alpha}(1+x)^{\frac{1}{2}\beta} P_n^{(\alpha,\beta)}(x)|^p \, dx \right\}^{1/p} \left\{ \int_0^1 (1+x)^{-\frac{1}{2}\beta p'} \, dx \right\}^{1/p'}.$$

Now $I \sim n^{-\frac{1}{2}}$ [17, p. 168], and the last integral is finite. Therefore, we may select an $A > 0$ such that

$$\left\{ \int_{-1}^{1} |(1-x)^{\frac{1}{2}\alpha}(1+x)^{\frac{1}{2}\beta} P_n^{(\alpha,\beta)}(x)|^p \, dx \right\}^{1/p} > A n^{-\frac{1}{2}} \tag{10.2}$$

for all n. From (10.1) and (10.2) we obtain the desired result.

THEOREM 10.2. *Let $w(x) \in \mathbf{B}$. Let $f(x) \in L_w^p$ for a value of p satisfying*

$$4 \max\left\{ \frac{\alpha+1}{2\alpha+3}, \frac{\beta+1}{2\beta+3} \right\} < p < 4 \min\left\{ \frac{\alpha+1}{2\alpha+1}, \frac{\beta+1}{2\beta+1} \right\},$$

and define $a_n = \int_{-1}^{1} f(x) p_n(x) w(x)\,dx$. Then

$$\lim_{N\to\infty} \int_{-1}^{1} |\, f(x) - \sum_{n=0}^{N} a_n p_n(x) \,|^p \, w(x)\,dx = 0.$$

Proof. According to a result of Pollard [13, p. 363], it suffices to prove that the kernels

$$K_{\pm}(x,t) = \left| \left\{ \left(\frac{1-t^2}{1-x^2} \right)^{\pm\frac{1}{4}} \left(\frac{w(t)}{w(x)} \right)^{\frac{1}{2}-1/p} - 1 \right\} / (x-t) \right| \tag{10.3}$$

have the property that $\int_{-1}^{1} K_{\pm}(x,t) g(t)\,dt$ belong to L^p whenever $g(t)$ does.

Our method of proof is an extension of one used already by him [14]. We shall deal only with K_+; a similar treatment may be used for K_-.

It suffices to show [14] that

$$F(x) = \int_{-1}^{1} K_+(x,t)\left(\frac{1-x^2}{1-t^2}\right)^{1/p} dt =$$

$$\int_{-1}^{1} \left| \left\{ \left(\frac{1-x}{1-t}\right)^a \left(\frac{1+x}{1+t}\right)^b \frac{T(t)}{T(x)} - \left(\frac{1-x}{1-t}\right)^c \left(\frac{1+x}{1+t}\right)^d \right\} \frac{1}{(x-t)} \right| dt$$

is bounded in $(-1,1)$, where

$$a = -\tfrac{1}{4} + \alpha(1/p - \tfrac{1}{2}) + 1/p, \qquad b = -\tfrac{1}{4} + \beta(1/p - \tfrac{1}{2}) + 1/p,$$

$$c = d = 1/p$$

and $T(x) = \{t(x)\}^{\frac{1}{2}-1/p}$. We remark that a, b, c, and d all lie in the open interval $(0,1)$.

Assume for the present that $0 \leqq x < 1$ and write $F(x)$ in the form

$$F(x) = \int_{-1}^{-\frac{1}{2}(1+x)} + \int_{-\frac{1}{2}(1+x)}^{x} + \int_{x}^{\frac{1}{2}(1+x)} + \int_{\frac{1}{2}(1+x)}^{1}$$

$= A(x) + B(x) + C(x) + D(x)$. (This decomposition is attributed by Pollard to P. Erdös.) The last integral is easily dealt with. Since the conditions on $t(x)$ imply that $T(x)$ is bounded above and below, $0 < m \leqq T(x) \leqq M < \infty$, we may write

$$D(x) \leqq K\{[(1-x)^a(1+x)^b/T(x)] \int_{\frac{1}{2}(1+x)}^{1} (1-t)^{-a} \mid x-t \mid^{-1} dt$$

$$+ (1-x)^c(1+x)^d \int_{\frac{1}{2}(1+x)}^{1} (1-t)^{-c} \mid x-t \mid^{-1} dt\}$$

$$\leqq K\{[(1-x)^a(1+x)^b/\tfrac{1}{2}m(1-x)] \int_{\frac{1}{2}(1+x)}^{1} (1-t)^{-a}\, dt$$

$$+ [(1-x)^c(1+x)^d/\tfrac{1}{2}(1-x)] \int_{\frac{1}{2}(1+x)}^{1} (1-t)^{-c}\, dt\} \leqq K.$$

The third integral is more difficult. Writing u for $t-x$, we get

$$(10.4) \quad C(x) = \int_{0}^{\frac{1}{2}(1-x)} u^{-1} \left| \left(\frac{1-x}{1-u-x}\right)^a \left(\frac{1+x}{1+u+x}\right)^b \frac{T(u+x)}{T(x)} \right.$$

$$\left. - \left(\frac{1-x}{1-u-x}\right)^c \left(\frac{1+x}{1+u+x}\right)^d \right| du.$$

By the mean value theorem,

$$T(u+x)/T(x) = 1 + uT'(x+\theta_x u)/T(x) \qquad (0 < \theta_x < 1)$$

for all x and u such that $-1 \leqq x \leqq 1$, $-1 \leqq u+x \leqq 1$. Formula (10.4) then gives us

$$C(x) \leqq \int_0^{\frac{1}{2}(1-x)} u^{-1} \left| \left(1+\frac{u}{1-u-x}\right)^a \left(1-\frac{u}{1+u+x}\right)^b - \left(1+\frac{u}{1-u-x}\right)^c \left(1-\frac{u}{1+u+x}\right)^d \right| du$$

$$+ \int_0^{\frac{1}{2}(1-x)} \left| \frac{T'(x+\theta_x u)}{T(x)} \left(1+\frac{u}{1-u-x}\right)^a \left(1-\frac{u}{1+u+x}\right)^b \right| du.$$

The first integral is known to be bounded [14]. In the second we observe that the conditions imposed upon $t(x)$ imply that $T'(x+\theta_x u)$ is continuous in $x+\theta_x u$, and hence bounded. Further, since $0 < a < 1$,

$$\left(1+\frac{u}{1-u-x}\right)^a \leqq 1 + \frac{u}{1-u-x}.$$

Hence

$$C(x) < K + K \int_0^{\frac{1}{2}(1-x)} (1-x)/(1-u-x)\,du < K.$$

The integrals $A(x)$ and $B(x)$ may be treated similarly, so that $F(x)$ is bounded for x in $[0, 1)$. The same arguments may clearly be used for x in $(-1, 0]$. This completes the proof of the theorem.

CORNELL UNIVERSITY.

BIBLIOGRAPHY.

[1] S. Banach, *Théorie des opérations linéaires*, Warsaw, 1932.

[2] S. Bernstein, "Sur les polynômes orthogonaux relatifs à un segment fini," *Journal de Mathematiques*, ser. (9), vol. 9 (1930), pp. 127-177; vol. 10 (1931), pp. 219-286.

[3] I. W. Busbridge, "A theory of general transforms for functions of the class $L^p(0, \infty)$ $(1 < p \leqq 2)$," *Quarterly Journal of Mathematics* (Oxford series), vol. 9 (1938), pp. 148-160.

[4] W. B. Caton and E. Hille, "Laguerre polynomials and Laplace integrals," *Duke Mathematical Journal*, vol. 12 (1945), pp. 217-242.

[5] U. Dini, *Serie de Fourier*, Pisa, 1880, pp. 190-277.

[6] G. H. Hardy, J. E. Littlewood, and G. Pólya, *Inequalities*, Cambridge, 1933.

[7] S. Kacmarz and H. Steinhaus, *Theorie der Orthogonalreihen*, Warsaw, 1935.

[8] H. Kober, "A note on approximation by rational functions," *Proceedings of the Edinburgh Mathematical Society*, vol. 7 (1946), pp. 123-133.

[9] T. M. MacRobert, "Asymptotic expressions for the Bessel functions and the Fourier-Bessel expansion," *ibid.*, vol. 39 (1921), pp. 13-20.

[10] R. E. A. C. Paley, "A remarkable system of orthogonal functions," *Proceedings of the London Mathematical Society*, vol. 34 (1932), pp. 241-279.

[11] H. Pollard, "The mean convergence of orthogonal series of polynomials," *Proceedings of the National Academy of Sciences U. S. A.*, vol. 32 (1946), pp. 8-10.

[12] ———, "The mean convergence of orthogonal series. I," *Transactions of the American Mathematical Society*, vol. 62 (1947), pp. 387-403.

[13] ———, "The mean convergence of orthogonal series. II," *ibid.*, vol. 63 (1948), pp. 355-367.

[14] ———, "The mean convergence of orthogonal series. III," *Duke Mathematical Journal*, vol. 16 (1949), pp. 189-191.

[15] M. Riesz, "Sur les fonctions conjugées," *Mathematische Zeitschrift*, vol. 27 (1927), pp. 218-244.

[16] J. Schauder, "Eine Eigenschaft der Haarschen Orthogonalsystem," *ibid.*, vol. 28 (1928), pp. 317-320.

[17] G. Szegö, *Orthogonal Polynomials*, American Mathematical Society Colloquium Publications, vol. 23, New York, 1939.

[18] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals*, Oxford, 1937.

[19] G. N. Watson, *A Treatise on the Theory of Bessel Functions*, Cambridge, 1944.

[20] E. T. Whittaker and G. N. Watson, *A Course in Modern Analysis*, Cambridge, 1947.

[21] W. H. Young, "On series of Bessel functions," *Proceedings of the London Mathematical Society*, ser. (2), vol. 18 (1920), pp. 163-200.

ON THE STRUCTURE OF SOME SUBGROUPS OF THE MODULAR GROUP.*

By EMIL GROSSWALD.

1. 1. Let us consider the group Γ of modular substitutions, where elements are the matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, where a, b, c, d are rational integers, satisfying

$$ad - bc = 1. \tag{1}$$

On account of the significance of a, b, c, d as coefficients of a linear fractional transformation, we consider the elements $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$ of Γ as identical. Imposing suitable supplementary conditions on a, b, c, d, we can define various subgroups of Γ. If, in particular, the supplementary condition is

$$c \equiv 0 \pmod{p}, \tag{2}$$

the corresponding matrices form a subgroup $\Gamma_0(p)$ of Γ, whose structure has been studied by H. Rademacher [4].[1] If, instead of (2),

$$b \equiv 0 \pmod{p} \tag{3}$$

holds, the results are perfectly analogous, as pointed out by H. Rademacher in the last paragraph of his above-mentioned paper.

In a similar way, we define the subgroups $\Gamma_0^0(p)$ by the conditions

$$b \equiv c \equiv 0 \pmod{p}; \tag{4}$$

and we define the "principal subgroups" $\Gamma(p)$ by the conditions

$$b \equiv c \equiv 0 \pmod{p}; \qquad a \equiv d \equiv 1 \pmod{p}. \tag{4'}$$

It is our purpose to study the structure of these subgroups, indicating for any given prime number p, the corresponding set of independent generators and independent defining relations for $\Gamma_0^0(p)$; hereby, the abstract structure of the corresponding subgroup will be completely defined. Moreover, though

* Received January 17, 1950.

[1] Numbers in square brackets refer to the bibliography, at the end of the paper.

the method does not permit to write directly the independent generators as functions of p, we can find, explicitly, the generators for every particular, numerical value of p.

1.2. The main results established in this paper are:

(a) The subgroups $\Gamma_0^0(2)$, $\Gamma_0^0(3)$ are generated respectively by 2, 3 independent generators; for $p > 3$, $\Gamma_0^0(p)$ is generated by

$$m = 2[(p+2)(p-1)/12] + 3$$

independent generators. The explicit form of these generators is found.

(b) For $p = 2$, $p = 3$, $\Gamma_0^0(2)$, $\Gamma_0^0(3)$ respectively are the free groups of 2, 3 generators respectively; for $p > 3$, the generators satisfy the following defining relations:

(i) Two of the generators satisfy

$$V^2(k_1', \lambda_1') = I; \qquad V^2(k_2', \lambda_2') = I \text{ if } p \equiv 1 \pmod 4. \tag{5}$$

(ii) Two of the generators satisfy

$$V^3(k_1'', \lambda_1'') = I; \qquad V^3(k_2'', \lambda_2'') = I \text{ if } p \equiv 1 \pmod 3. \tag{6}$$

(iii) Two of the generators satisfy relations like (5), and two satisfy relations like (6), if $p \equiv 1 \pmod{12}$.

(iv) No defining relations exist if $p \equiv -1 \pmod{12}$; $\Gamma_0^0(p)$ is then the free group of m independent generators.

(c) For $p = 2$, both independent generators of $\Gamma_0^0(2)$ are parabolic; for $p \geqq 3$, three of the independent generators of $\Gamma_0^0(p)$ are parabolic.

(d) The genus of a fundamental region corresponding to $\Gamma_0^0(2)$ and $\Gamma_0^0(3)$ is zero; for $p > 3$, the genus is given by

$$g = [(p+2)(p-1)/12] - 1 + r(r^2-25)/24,$$

where $p \equiv r \pmod{12}$; $-5 \leqq r \leqq +5$.

(e)[2] If $p = 2$, the principal subgroup $\Gamma(2)$ is identical with the subgroup $\Gamma_0^0(2)$; for $p = 3$, $\Gamma(3)$ is the free group of 3 generators; for $p > 3$, and $p \equiv -1 \pmod{12}$, $\Gamma(p)$ is a free group of $p(p^2-1)/12 + 1$ generators. The method follows closely that of Professor H. Rademacher [4].

---

[2] After the present paper had been completed, it was brought to our attention that the structure of the principal subgroups had been completely determined by H. Frasch (see [7]) who proved the results from (e) without the restriction $p \equiv -1 \pmod{12}$.

Results of K. Reidemeister [5] and O. Schreier [6] have also been used. We wish to take this opportunity to express our gratitude to Professor H. Rademacher, who suggested the subject, and whose methods and advice have made this paper possible.

2. 1. If H is a subgroup of the group G, we may write G as the sum of H and its co-sets, as

$$G = H + HR_1 + HR_2 + \cdots + HR_n + \cdots, \tag{7}$$

where

$$R_n \varepsilon G, \quad R_n \notin H \qquad n = 1, 2, \cdots. \tag{8}$$

With $R_0 = I$, we may write (7) as

$$G = \sum_{n=0}^{N} H \cdot R_n \qquad N \leqq +\infty. \tag{7'}$$

The elements of the set $\{R_n\}$, including $R_0 = I$, are called the "representatives" of the decomposition (7'). Given any element $M \varepsilon G$, let HR_m be the co-set to which it belongs; then, following Schreier, we shall wirte $\bar{M}$ for the corresponding representative R_m. K. Reidemeister [5] has shown that if $\{S_i\}$ $(i = 1, 2, \cdots)$ are the generators of G, then all generators of H are found, eliminating from the expressions

$$R_n S_i \overline{R_n S_i}^{-1} \tag{9}$$

those, which reduce to the identity, when, in (9), the R_n and S_i range independently over all representatives $\{R_n\}$ of the decomposition (7'), respectively over all generators $\{S_i\}$ of G.

In order to be able to apply the results of O. Schreier [6] we have to take care that the system of representatives $\{R_n\}$ be chosen in such a way, as to satisfy Schreier's condition:[3] "Whenever $R_n = \prod_{i=1}^{n} S_i^{\varepsilon_i}$ is a representative of its co-set, also all sections of the product, $\prod_{i=1}^{j} S_i^{\varepsilon_i}$ $(j = 1, 2, \cdots, n-1)$ shall be representatives of their co-sets." Once this condition is satisfied, all defining relations of H are found (see [6] and [4]) by

$$R_n p(S_i) R_n^{-1} = I, \tag{10}$$

where the R_n range over all elements of $\{R_n\}$ (including $R_0 = I$) and

[3] See [6], page 177, condition "F."

$p(S_i) = I$ are the defining relations of G, and where we express the left-hand member by the generators (9) of H.

2. 2. On account of the preceding results, we shall proceed in the following order:

1) A system of representatives of the decomposition

$$\Gamma = \sum_{n=0}^{N} \Gamma_0^0(p) R_n \tag{11}$$

is defined, satisfying Schreier's condition.

2) By (9), the generators of $\Gamma_0^0(p)$ are found, using the known generators of Γ:[4]

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{12}$$

3) By (10), the defining relations of $\Gamma_0^0(p)$ are determined, using the well-known fact,[4] that the defining relations of Γ are:

$$T^2 = I \quad \text{and} \quad (TS)^3 = I. \tag{13}$$

4) Use is made of the defining relations, in order to eliminate as many generators as possible, until only independent generators and independent defining relations are left.

5) The number of independent generators is then found as a function of p, and also the form and number of the defining relations established in their dependence on p.

6) As a consequence of preceding results, a relation will be found, relating directly the genus of a fundamental region, corresponding to the subgroup $\Gamma_0^0(p)$, to the prime number p.

3. The representatives. Let

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \varepsilon\, \Gamma_0^0(p) \quad \text{and} \tag{14}$$

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = R, \qquad \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} R \varepsilon\, \Gamma. \tag{15}$$

Then, from

[4] See, e. g. [4].

$$(16) \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} a\alpha + b\gamma & a\beta + b\delta \\ c\alpha + d\gamma & c\beta + d\delta \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix},$$

using (14) and (4), we obtain:[5]

$$(17) \qquad a\alpha \equiv A \qquad a\beta \equiv B \qquad d\gamma \equiv C \qquad d\delta \equiv D.$$

By simple, algebraic computations, using (17), we find the following results:

3.1. If $A \equiv 0$ (implying, by (1) and (15), that $BC \equiv -1$, so that $B \not\equiv 0$, $C \not\equiv 0$), then (16) can be satisfied with $R = TS^k = \begin{pmatrix} 0 & -1 \\ 1 & k \end{pmatrix}$, where k has to satisfy

$$(18) \qquad D \equiv Ck.$$

Adding to (18) the condition

$$(19) \qquad 0 \leqq k \leqq p-1,$$

the value of k is uniquely determined. We remark, that for $D \equiv 0$, $k \equiv 0$ and $R = T$; for $D \not\equiv 0$,

$$(19') \qquad 1 \leqq k \leqq p-1.$$

Once k is determined by (18) and (19), or (19′), we find the corresponding matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} R^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$
$$\times S^{-k}T^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} k & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} Ak - B & A \\ Ck - D & C \end{pmatrix}.$$

On account of $A \equiv 0$, we check, by (18), that actually $b = A \equiv 0$ $c = Ck - D \equiv 0$, so that indeed: $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \varepsilon\, \Gamma_0^0(p)$, as it had to be.

3.2. In an entirely similar way, still using (17), we find, if $B \equiv 0$: $R = TS^kT = \begin{pmatrix} 1 & 0 \\ -k & 1 \end{pmatrix}$ with k satisfying

$$(20) \qquad kD + C \equiv 0 \qquad 1 \leqq k \leqq p-1,$$

---

[5] Unless otherwise specified, all congruences have to be understood (mod p). For simplicity, we shall consistently suppress these words, except where confusion could arise.

and, as before, once k has been found by (20), the corresponding matrix[6]

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is found by } \begin{pmatrix} A & B \\ C & D \end{pmatrix} R^{-1} = \begin{pmatrix} A+Bk & B \\ C+Dk & D \end{pmatrix} \equiv \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix} \varepsilon\, \Gamma_0^0(p).$$

3.3. Similarly, if $C \equiv 0$ (implying $A \not\equiv 0$, $D \not\equiv 0$), then, by (17), $R = S^l = \begin{pmatrix} 1 & l \\ 0 & 1 \end{pmatrix}$, where l satisfies

$$Al \equiv B, \qquad 0 \leqq l \leqq p-1. \tag{21}$$

If, in particular, also $B \equiv 0$, then $R = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, since $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ then belongs to $\Gamma_0^0(p)$. Excluding this case, l has to satisfy

$$1 \leqq l \leqq p-1 \tag{21'}$$

Once l determined by (21) or (21'); the corresponding matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is found as before.

3.4. Finally, in the general case, when $A \not\equiv 0$, $B \not\equiv 0$, $C \not\equiv 0$, we find, still using (17), that $R = TS^kTS^l = \begin{pmatrix} 1 & l \\ -k & 1-kl \end{pmatrix}$, where k and l are determined by

$$k \equiv -AC, \qquad 1 \leqq k \leqq p-1, \tag{22}$$

$$Al \equiv B, \qquad 1 \leqq l \leqq p-1. \tag{23}$$

The last relation of (17) shows also that

$$kl \equiv 1 \tag{24}$$

if and only if $D \equiv 0$.

All possible cases are now exhausted and we arrange, conveniently, the found representatives into the following array (I), which shows that they satisfy Schreier's condition:

(I)

S–I–T–TS	–	TS^2 –· · · · –		TS^{p-1}
S^2		TST	TS^2T	$TS^{p-1}T$
S^3		$TSTS$	TS^2TS	$TS^{p-1}TS$
S^4		$TSTS^2$	TS^2TS^2	$TS^{p-1}TS^2$
⋮		⋮	⋮	⋮
S^{p-1}		$TSTS^{p-1}$	TS^2TS^{p-1}	$TS^{p-1}TS^{p-1}$.

[6] By $\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \equiv \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$ we understand $a_{ij} \equiv b_{ij}$ for all i and j.

4.1. The generators of $\Gamma_0^0(p)$. From (9) we obtain, on account of (12) the generators

$$(25) \qquad RS\,\overline{RS}^{-1} \qquad (26) \qquad RT\,\overline{RT}^{-1}.$$

where R takes successively all values of the array (I). Using (25), we obtain for $R = I$.

$$IS\,\overline{IS}^{-1} = S\,S^{-1} = S\,S^{-1} = I.$$

Similarly, as simple computations show, we obtain again the identity, if we take:

$$\begin{aligned} &R = S^k && 1 \leqq k \leqq p-2 \\ &R = TS^k && 0 \leqq k \leqq p-2 \\ &R = TS^kT && 1 \leqq k \leqq p-1 \\ &R = TS^kTS^l && 1 \leqq k \leqq p-1;\quad 0 \leqq l \leqq p-2. \end{aligned}$$

There remain, therefore, still to be considered, only the cases

(i) $R = S^{p-1}$: $RS = S^p \in \Gamma_0^0(p)$, so that $\overline{RS} = I$ and

$$(27) \qquad RS\,\overline{RS}^{-1} = S^p\,I^{-1} = S^p;$$

(ii) $R = TS^{p-1}$: $RS = TS^p = \begin{pmatrix} 0 & -1 \\ 1 & p \end{pmatrix}$, so that, by (18), (19), $\overline{RS} = T$, and

$$(28) \qquad RS\,\overline{RS}^{-1} = TS^pT^{-1};$$

(iii) $R = TS^kTS^{p-1}$: $RS = TS^kTS^p = \begin{pmatrix} 1 & p \\ -k & 1-kp \end{pmatrix}$, so that, by (20), $\overline{RS} = TS^{k'}T$, where k' has to satisfy $k'(1-kp) - k \equiv 0$, $1 \leqq k' \leqq p-1$, wherefrom, using the second relation from (20), $k' = k$, so that $\overline{RS} = TS^kT$ and

$$(29) \qquad RS\,\overline{RS}^{-1} = TS^kTS^pT^{-1}S^{-k}T^{-1}.$$

4.2. Let us consider now the expression (26). For

$$\begin{aligned} &R = I, \\ &R = TS^k, && 0 \leqq k \leqq p-1 \\ &R = TS^kT, && 1 \leqq k \leqq p-1, \end{aligned}$$

$RT = \overline{RT}$, so that $RT\,\overline{RT}^{-1} = I$. We have still to consider:

(i) $R = S^k$: $RT = S^kT = \begin{pmatrix} k & -1 \\ 1 & 0 \end{pmatrix}$, so that, by (22), (23), (24),

$$\overline{RT} = TS^{k'}TS^{l'} \text{ with}$$

$$k' \equiv -k, \qquad 1 \leqq k' \leqq p-1 \tag{30}$$
$$kl' \equiv -1, \qquad 1 \leqq l' \leqq p-1.$$

Let us consistently mark by k_*, the number x, satisfying the two relations (32):

$$kx \equiv -1, \qquad 1 \leqq x \leqq p-1. \tag{32}$$

Then we obtain from (30), (31):

$$l' = k_*,\ k' = p-k \text{ so that } \overline{RT} = TS^{p-k}TS^{k_*} \text{ and} \tag{33}$$

$$RT\,\overline{RT}^{-1} = S^kTS^{-k_*}T^{-1}S^{-p+k}T^{-1}. \tag{34}$$

If $R = TS^kTS^l$, then $RT = TS^kTS^lT = \begin{pmatrix} l & -1 \\ 1-kl & k \end{pmatrix}$ and two cases arise:

(ii) let $1 - kl \equiv 0$; then, by **3. 3**, $\overline{RT} = S^{l'}$ with $ll' \equiv -1$, $1 \leqq l' \leqq p-1$, so that, by (32), $l' = l_*$ and $RT = S^{l_*}$. Then, as $kl \equiv 1$, $(-k)$, $l \equiv -1$, $k \equiv -l_*$, $k = p - l_*$, so that

$$RT\,\overline{RT}^{-1} = TS^{p-l_*}TS^lT \cdot S^{-l_*}. \tag{35}$$

(iii) Let $1 - kl \equiv 0$. Then, by **3. 4**:

$$\overline{RT} = TS^{k'}TS^{l'}, \text{ with } ll' \equiv -1,\ 1 \leqq l' \leqq p-1,$$
$$k' \equiv -l(1-kl) \equiv l(kl-1),\ 1 \leqq k' \leqq p-1.$$

Given any two numbers k and l, $1 - kl \not\equiv 0$, we shall write consistently k_l for the number x satisfying

$$x \equiv l(kl-1) \qquad 1 \leqq x \leqq p-1; \tag{36}$$

therefore, we can write using also (32),

$$l' = l_*,\ \ k' = k_l, \text{ and } \overline{RT} = TS^{k_l}TS^{l_*}, \text{ so that:}$$

$$RT\,\overline{RT}^{-1} = TS^kTS^lTS^{-l_*}T^{-1}S^{-k_l}T^{-1}. \tag{37}$$

We shall designate, conveniently, the generators of $\Gamma_0{}^0(p)$ by capital letters,[7]

[7] In the choice of these letters, we have been guided by the endeavour to match, as far as possible, the designations of H. Rademacher, [4] in order to permit an easy comparison of the results.

marking also the parameters on which they depend in each case. The results from (28)-(37) can, then, be condensed into the following array (II):

$$
\begin{array}{lll}
Z = S^p & & \\
U = TS^pT^{-1} & & \\
X(k) = TS^kTS^pT^{-1}S^{-k}T^{-1} & & 1 \leqq k \leqq p-1 \\
Y(k) = S^kTS^{-k_*}T^{-1}S^{-p+k}T^{-1} & & 1 \leqq k \leqq p-1 \\
W(k) = TS^{p-k_*}TS^kTS^{-k_*} & & 1 \leqq k \leqq p-1 \\
V(k,l) = TS^kTS^lTS^{-l_*}T^{-1}S^{-k_1}T^{-1} & & \begin{cases} 1 \leqq k \leqq p-1 \\ 1 \leqq l \leqq p-1 \\ kl \not\equiv 1. \end{cases}
\end{array}
\tag{II}
$$

We may remark at once, that the generators $V(k,l)$ exist only for $p \geqq 3$, as for $p = 2$, $k = l = 1$ and $kl = 1$.

5. 1. The defining relations. The relations (10) become, by (13),

$$RT^2R^{-1} = I, \quad \text{and} \quad R(TS)^3R^{-1} = I, \tag{38, 39}$$

where R ranges over all representatives of the array (I) and the left-hand sides have to be expressed as products of the generators (II). Let us consider first (38). For $R = I$ and $R = T$ we obtain, after simple computations, the trivial result $U \cdot U^{-1} = I$. Similarly, for $R = TS^k$ and $R = TS^kT$, (38) reduces to the identity $X(k) \cdot X^{-1}(k) = I$. For $R = TS^kTS^l$, however, we obtain successively:

$$
\begin{aligned}
TS^kTS^l \cdot T^2 \cdot S^{-l}T^{-1}S^{-k}T^{-1} &= TS^kTS^lTS^{-l_*}T^{-1}S^{-k_1}T^{-1} \\
&\quad \cdot TS^{k_1}TS^{l_*}TS^{-l}T^{-1}S^{-k}T^{-1} = I.
\end{aligned}
$$

$$V(k,l) \cdot TS^{k_1}TS^{l_*}TS^{-l}T^{-1}S^{-k}T^{-1} = I. \tag{40}$$

By (36), with $k = k_1$ and $l = l_*$, we obtain successively:

$$
\begin{aligned}
(k_1)_1 \equiv l_*(k_1l_* - 1) &\equiv l_*\{l(kl-1)l_* - 1\} \\
&\equiv l_*(-kl + 1 - 1) \equiv -kll_* \equiv k.
\end{aligned}
$$

As, obviously, also $(l_*)_* = l$, it follows that

$$TS^{k_1}TS^{l_*}TS^{-l}T^{-1}S^{-k}T^{-1} = TS^{k_1}TS^{l_*}TS^{-l_{**}}T^{-1}S^{-(k_1)_1}T^{-1} = V(k_1, l_*),$$

by (II), and (40) becomes

$$V(k,l) \cdot V(k_1, l_*) = I. \tag{41}$$

5.2. With $R = S^k$ we obtain, from (38)

$$S^kT^2S^{-k} = S^kT \cdot S^{-k_\bullet}T^{-1}S^{-p+k}T^{-1} \cdot TS^{p-k}TS^{k_\bullet} \cdot TS^{-k} = I, \tag{42}$$

or, by (II):

$$Y(k) \cdot W(k_\bullet) = I, \tag{42}$$

and that is all we obtain from (38).

5.3. From (39) we obtain for

$$R = I : I(TS)^3I^{-1} = TSTSTS = TSTSTS^{-p+1} \cdot S^p = I,$$

so that, by (II)

$$W(1) \cdot Z = I. \tag{43}$$

$$R = T : T(TS)^3T^{-1} = T^2STSTST^{-1} = TS^pT^{-1} \cdot TS^{-p}T \cdot STSTST^{-1}$$
$$= U \cdot U^{-1} \cdot STSTST^{-1} = STS^{-p+1}T^{-1}S^{-p+1}T^{-1} \cdot TS^{p-1}TS^pTST^{-1}$$
$$= Y(1) \cdot TS^{p-1}TS^pT^{-1}S^{-p+1}T^{-1} \cdot U = I,$$

by repeated use of the definitions (II), or, finally,

$$Y(1) \cdot X(p-1) \cdot U = I; \tag{44}$$

$R = TS$ leads again to (43). $R = TS^k$, $2 \leqq k \leqq p-1$, leads to

$$TS^k \cdot TSTSTS \cdot S^{-k}T^{-1} = TS^kTSTSTS^{-k+1}T^{-1} = TS^kTSTS^{-p+1}T^{-1}S^{-k_1}T^{-1}$$
$$\cdot TS^{k_1}TS^pTS^{-k+1}T^{-1} = V(k,1) \cdot TS^{k_1}TS^pTS^{-k+1}T^{-1} = I.$$

Observing that, with $l = 1$, by (36), $k_1 = l(kl-1) = k-1$, the last equality becomes, by (II),

$$V(k,1) \cdot X(k-1) = I \qquad 2 \leqq k \leqq p-1. \tag{45}$$

$R = TS^kT$, $1 \leqq k \leqq p-2$ leads to

$TS^kT \cdot (TS)^3 \cdot T^{-1}S^{-k}T^{-1} = TS^{k+1}TSTST^{-1}S^{-k}T^{-1} = I$, whence

$V(k+1,1) \cdot X(k) = I$, $1 \leqq k \leqq p-2$, same as (45). Also, for $k = p-1$, $R = TS^{p-1}T$ leads to nothing new, as (39) becomes $U \cdot Y(1) \cdot X(p-1) = I$, not essentially distinct from (44).

$R = TS^kTS$ leads again to (45).

5.4. $R = TS^kTS^l$ with k and l satisfying

$$2 \leqq l \leqq p-1; \qquad 1 \leqq k \leqq p-1, \tag{46}$$

leads to the following three separate cases:

5.4.1. If

(47) $kl \equiv 1$ then, by (32) and (46), $l = p - k_*$ and (39) becomes

$$\begin{aligned} TS^kTS^{p-k_*} \cdot TSTSTS \cdot S^{-p+k_*}T^{-1}S^{-k}T^{-1} \\ = TS^kTS^{p-k_*}T^{-1}S^{-p+k} \cdot S^{p-k+1}TSTS^{-p+k_*+1}T^{-1}S^{-k}T^{-1} \\ = Y^{-1}(p-k) \cdot S^{p-k+1}TS^{-(p-k+1)_*}T^{-1}S^{-k+1}T^{-1} \\ \cdot TS^{k-1}TS^{1+(p-k+1)_*}TS^{-p+k_*+1}T^{-1}S^{-k}T^{-1} = I, \end{aligned}$$

whence, by (II),

(48) $$Y^{-1}(p-k) \cdot Y(p-k+1) \cdot V(k-1 \cdot \bar{l}) = I, \quad 2 \leqq k \leqq p-1,$$

with

(49) $$\bar{l} \equiv 1 + (p-k+1)_*, \quad 1 \leqq \bar{l} \leqq p-1.$$

In proving (48) and (49), we have used

(50) $$\bar{l}_* \equiv \{1 + (p-k+1)_*\}_* \equiv (p-k_*-1), \qquad \text{and}$$

(51) $$(k-1)_1 = k,$$

whose proofs follow.

Proof of (50). Writing x for the product:

$$x = \{1 + (p-k+1)_*\}(p-k_*-1),$$

(50) is equivalent to the congruence: $x \equiv -1$; multiplying x by $(p-k+1)$, we obtain successively:

$$\begin{aligned} x(p-k+1) &\equiv \{p-k+1-1\}(p-k_*-1) \equiv -k(-k_*-1) \\ &\equiv -1+k \equiv -(p-k+1), \end{aligned}$$

wherefrom $x \equiv -1$ follows, proving (50).

Proof of (51). By straightforward computation, and using the definition (36), we obtain successively:

$$\begin{aligned} (k-1)_1 &\equiv \{1 + (p-k+1)_*\}\{(k-1)(1 + (p-k+1)_*) - 1\} \\ &\equiv \{1 + (p-k+1)_*\}\{-(p-k+1)(1 + (p-k+1)_*) - 1\} \\ &\equiv \{1 + (p-k+1)_*\}\{-(p-k+1) + 1 - 1\} \\ &\equiv -(p-k+1)\{1 + (p-k+1)_*\} \\ &\equiv p + k - 1 + 1 \equiv k, \end{aligned}$$

finishing the proof of (48).

5.4.2. If

$$k(l-1) \equiv 1, \tag{52}$$

then $kl - k \equiv 1$, or, multiplying by $k_\bullet$: $-l+1 \equiv k_\bullet$, $l \equiv -k_\bullet + 1$, so that [8]

$$l = p + 1 - k_\bullet,$$

and (39) becomes

$$\begin{aligned} &TS^kTS^l \cdot TSTSTS \cdot S^{-l}T^{-1}S^{-k}T^{-1} \\ &\quad = TS^kTS^lTS^{-l_\bullet}T^{-1}S^{-k_1}T^{-1} \cdot TS^{k_1}TS^{l_\bullet+1}TSTS^{-l+1}T^{-1}S^{-k}T^{-1} \\ &\quad = V(k,l) \cdot TS^{k_1}TS^{l_\bullet} \cdot STSTS \cdot S^{-l}T^{-1}S^{-k}T^{-1} = I. \end{aligned}$$

As

$$\begin{aligned} k_1 &\equiv l(kl-1) \equiv (p+1-k_\bullet)(k(p+1-k_\bullet)-1) \\ &\equiv (1-k_\bullet)(k+1-1) \equiv k+1, \end{aligned}$$

and as $STSTS = T^{-1} = T$, the above relation becomes

$$V(k,l) \cdot TS^{k_1}TS^{l_\bullet}TS^{-l}T^{-1}S^{-k}T^{-1} = V(k,l) \cdot V(k_1, l_\bullet) = I,$$

because

$$\begin{aligned} (k_1)_1 &\equiv l_\bullet(k_1 l_\bullet - 1) \equiv l_\bullet\{l(kl-1)l_\bullet - 1\} \equiv -(kl-1)l_\bullet - l_\bullet \\ &\equiv k + l_\bullet - l_\bullet \equiv k. \end{aligned}$$

The relation found is only a particular case (corresponding to $l = p + 1 - k_\bullet$), of that already obtained under (41).

5.4.3. Let now

$$kl \not\equiv 1 \quad \textit{and} \quad k(l-1) \not\equiv 1. \tag{53}$$

Then (39) becomes

$$TS^kTS^l \cdot TSTSTS \cdot S^{-l}T^{-1}S^{-k}T^{-1} = I,$$

whence, using the notations

$$l_{j+1} \equiv (l_j)_\bullet + 1, \qquad l = l_0,\ 1 \leqq l_j \leqq p-1, \tag{54}$$

$$k_{j+1} \equiv l_j(k_j l_j - 1), \qquad k = k_0,\ 1 \leqq k_j \leqq p-1. \tag{55}$$

we obtain successively

---

[8] Notice that $k_\bullet \neq 1$, as $k_\bullet = 1$ implies $k = p-1$, whence, by (52), $l - 1 \equiv p - 1$ contradicting (46).

$$TS^kTS^lTS^{-l_\bullet}T^{-1}S^{-k_1}T^{-1}\cdot TS^{k_1}TS^{l_\bullet+1}TSTS^{-l+1}T^{-1}S^{-k}T^{-1}$$
$$= V(k,l)\cdot TS^{k_1}TS^{l_1}TS^{-l_{1\bullet}}T^{-1}S^{-(k_1)_1}T^{-1}\cdot TS^{(k_1)_1}TS^{l_{1\bullet}+1}TS^{-l+1}T^{-1}S^{-k}T^{-1}$$
$$= V(k,l)\cdot V(k_1,l_1)\cdot TS^{k_2}TS^{l_2}TS^{-l+1}T^{-1}S^{-k}T^{-1} = I,$$

or

$$V(k,l)\cdot V(k_1,l_1)\cdot V(k_2,l_2) = I. \tag{56}$$

In order to justify the last equality, we have to show that:

$$l_{2\bullet} = l-1, \text{ respectively } l_3 = l_{2\bullet}+1 = l \tag{57}$$

and

$$(k_2)_1 = k_3 = k. \tag{58}$$

Proof of (57). We have to prove that $l_2(l-1)\equiv -1$, or, by the definition of l_2,

$$(l_{1\bullet}+1)(l-1)\equiv -1,$$

leading to $l_{1\bullet}l - l_{1\bullet}+l\equiv 0$, and, successively, upon multiplication by l_1:

$$-l+1+ll_1\equiv 0;\quad l(l_1-1)\equiv -1;\quad l(l_\bullet+1-1)\equiv -1,$$

which is true. As all steps are reversible, (57) holds.

Proof of (58). Using repeatedly the definitions (54), (55), we obtain (58) by straightforward computation:

$$k_3 = (k_2)_1 \equiv l_2(k_2l_2-1)\equiv (1+l_{1\bullet})\{l_1(k_1l_1-1)(1+l_{1\bullet})-1\}$$
$$\equiv (1+l_{1\bullet})\{l_1(k_1l_1-1)-k_1l_1+1-1\}\equiv (1+l_{1\bullet})\{k_1l_1(l_1-1)-l_1\}$$
$$\equiv (1+l_{1\bullet})(k_1l_1l_\bullet - l_1)\equiv k_1(l_\bullet+1)l_\bullet - l_1 - k_1l_\bullet+1$$
$$\equiv k_1l_\bullet l_\bullet - l_1+1\equiv l_\bullet l(kl-1)l_\bullet - (l_\bullet+1)+1$$
$$\equiv -(kl-1)l_\bullet - l_\bullet\equiv k, \text{ i. e. (58).}$$

From (57), (58) it results also that, starting with a pair of numbers $(k,l) = (k_0,l_0)$, satisfying the conditions (46) and (53), the sucessive, iterated pairs of numbers have period 3, i. e. $k_{n+3} = k_n$, $l_{n+3} = l_n$. Therefore, starting, e. g. in (56), with $V(k_1,l_1)$, the corresponding relation will be:

$$V(k_1,l_1)\cdot V(k_2,l_2)\cdot V(k_3,l_3) = V(k_1,l_1)\cdot V(k_2,l_2)\cdot V(k,l) = I,$$

which is precisely equivalent to (56). We shall use this remark, when we count the number of independent relations.

5. 5. We have, finally, to consider $R = S^k$, and obtain from (39):

$$R(TS)^3R^{-1} = S^kTSTSTS\cdot S^{-k} = S^kTSTS^{-p+1}T^{-1}S^{-k_1}T^{-1}\cdot TS^{k_1}TS^pTS^{1-k}$$
$$= V(k,1)\cdot TS^{k-1}TS^pTS^{-(k-1)} = V(k,1)\cdot X(k-1) = I,$$

which is valid for $2 \leqq k \leqq p-1$. This contains, however, nothing new, cf. (45). For $k=1$, the relation reduces also to one already found, namely:

$$STSTST = Y(1) \cdot X(p-1) \cdot U = I, \text{ i. e. (44).}$$

5. 6. The complete list of defining relations is, therefore:

$$W(1) \cdot Z = I$$

$$Y(1) \cdot X(p-1) \cdot U = I$$

$$V(k,l) \cdot X(k-1) = I \qquad 2 \leqq k \leqq p-1$$

(III) $$Y^{-1}(p-k) \cdot Y(p-k+1) \cdot V(k-1, \bar{l}) = I$$
$$2 \leqq k \leqq p-1, \bar{l} = 1 + (p-k+1).$$

$$Y(k) \cdot W(k_\bullet) = I \qquad 1 \leqq k \leqq p-1$$

$$V(k,l) \cdot V(k_1, l_\bullet) = I \qquad 1 \leqq k \leqq p-1, kl \not\equiv 1$$
$$1 \leqq l \leqq p-1$$

$$V(k,l) \cdot V(k_1, l_1) \cdot V(k_2, l_2) = I$$
$$1 \leqq k \leqq p-1, kl \not\equiv 1$$
$$2 \leqq l \leqq p-1, k(l-1) \not\equiv 1.$$

As already pointed out,[9] the generators $V(k,l)$ do not exist for $p=2$, and (III) reduces to the three equations:

$$W(1) \cdot Z = I \qquad Y(1) \cdot X(1) \cdot U = I \qquad Y(1) \cdot W(1) = I.$$

For $p=3$, there are only two generators $V(k,l)$, namely $V(1,2)$ and $V(2,1)$, and in (III) the last relations do not exist.[10] It is easy to see that in this case there exists exactly one relation of each type, except the last (which we saw, does not exist at all) and

$Y(k) \cdot W(k_\bullet) = I$, which gives two relations, as k may be 1 or 2.

For $p > 3$, all relations of (III) exist effectively. This is obvious for all relations, excepting, on the one hand, (48), for which case we have to prove that $(k-1) \cdot \bar{l} \not\equiv 1$ for $2 \leqq k \leqq p-1$, on the other hand, (56), for which case we have to show that $k_1 l_1 \not\equiv 1$, $l_1 \not\equiv 1$ and $k_2 l_2 \not\equiv 1$, $l_2 \not\equiv 1$, whenever k and l satisfy (46) and (53).

---

[9] See end of paragraph **4. 2.**

[10] Because, with $l=2$, and $k=1$, or $k=2$, either $kl=1$, or $k(l-1)=1$; see also **§ 7. 2** particularly (69), which admits a single solution and (70) which, on account of (71) admits no solution.

For the case (48), from $\bar{l} = 1 + (p-k+1)\cdot$, with $2 \leqq k \leqq p-1$, (hence $k \not\equiv 1$), we have

$$(k-1)\cdot\bar{l} \equiv -(p-k+1)\{1+(p-k+1)\cdot\}$$
$$\equiv -(p-k+1)+1 \equiv k \not\equiv 1,$$

proving the existence of (48).

For (56), assume, contrary to the assertion, that $k_1 l_1 \equiv 1$. Substituting for k_1 and l_1 their values from (54), (55), we obtain

$l(kl-1)(l\cdot+1) \equiv 1$, whence $kl(l-1) \equiv l$ and, as $l \not\equiv 0$, $k(l-1) \equiv 1$, contradicting (53); proving, therefore, that $k_1 l_1 \not\equiv 1$. If now, contrary to our assertion, $k_2 l_2 \equiv 1$, it follows, as before, that

$l_1(k_1 l_1 - 1)(l_1\cdot + 1) \equiv 1$, whence $k_1 l_1 (l_1 - 1) \equiv l_1$, so that either $l_1 \equiv 0$, or $k_1(l_1-1) \equiv 1$. If $l_1 \equiv 0$, then $l\cdot + 1 \equiv 0$, $l \equiv 1$, contrary to (46). If $k_1(l_1-1) \equiv 1$, then $l(kl-1)l\cdot \equiv 1$, whence $kl \equiv 0$ would follow, which is impossible by (46).[11]

Obviously, also $l_1 \equiv l\cdot + 1 \not\equiv 1$, as $l\cdot \not\equiv 0$, and similarly for l_2. That finishes the proof of our assertion, that for $p > 3$, all relations (III) exist effectively.

6. Reduction to a system of independent generators and defining relations.

6.1. For $p = 2$, (II) and (III) reduce, respectively, to

$$\text{(II')}\quad \begin{aligned} Z &= S^2 \\ U &= TS^2T^{-1} \\ X(1) &= TSTS^2T^{-1}S^{-1}T^{-1} \\ Y(1) &= STS^{-1}T^{-1}S^{-1}T^{-1} \\ W(1) &= TSTSTS^{-1}; \end{aligned} \qquad \text{(III')}\quad \begin{aligned} W(1)\cdot Z &= I \\ Y(1)\cdot X(1)\cdot U &= I \\ Y(1)\cdot W(1) &= I. \end{aligned}$$

The three relations (III') permit us to eliminate three of the five generators (II'), expressing them with the help of the other two, e. g.:

$$W(1) = Z^{-1}, \qquad Y(1) = Z, \qquad X(1) = Y^{-1}(1)\cdot U^{-1} = Z^{-1}\cdot U^{-1}.$$

---

[11] This conclusion would not hold, if the congruences were taken modulo n, with n composite. Difficulties of this sort make complicated an extension of the theory to the subgroups $\Gamma_0^0(n)$.

We remain with the two independent generators:

$$Z = S^2 = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \text{ and } U = TS^2T^{-1} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix},$$

and no defining relations. Therefore $\Gamma_0^0(2)$ *is the free group of two generators.*

6. 2. For $p = 3$, (II) and (III) become, respectively:

$$\begin{aligned}
&Z = S^3, \\
&U = TS^3T^{-1}, \\
&X(1) = TSTS^3T^{-1}S^{-1}T^{-1} && W(1) \cdot Z = I; \\
&X(2) = TS^2TS^3T^{-1}S^{-2}T^{-1} && Y(1) \cdot X(2) \cdot U = I. \\
&Y(1) = STS^{-2}T^{-1}S^{-2}T^{-1} && V(2,1) \cdot X(1) = I \\
\text{(II}'')\quad &Y(2) = S^2TS^{-1}T^{-1}S^{-1}T^{-1} && Y^{-1}(1) \cdot Y(2) \cdot V(1,2) = I \\
&W(1) = TSTSTS^{-2} \qquad \text{(III}'') && Y(1) \cdot W(2) = I \\
&W(2) = TS^2TS^2TS^{-1} && Y(2) \cdot W(1) = I \\
&V(1,2) = TSTS^2TS^{-1}T^{-1}S^{-2}T^{-1} && V(1,2) \cdot V(2,1) = I. \\
&V(2,1) = TS^2TSTS^{-2}T^{-1}S^{-1}T^{-1},
\end{aligned}$$

The seven relations (III′) permit us to eliminate seven of the ten generators (II″) as follows:

By the first of (III″) :		$W(1) = Z^{-1}$
" " 6-th " " :		$Y(2) = W^{-1}(1) = Z$
" " 4-th " " :		$Y(1) = Y(2) \cdot V(1,2) = Z \cdot V(1,2)$
" " 5-th " " :		$W(2) = Y^{-1}(1) = V^{-1}(1,2) \cdot Z^{-1}$
" " 2-th " " :		$X(2) = Y^{-1}(1) \cdot U^{-1} = V^{-1}(1,2) \cdot Z^{-1} \cdot U^{-1}$
" " 3-th " " :		$X(1) = V^{-1}(2,1) = V(1,2)$, because,
" " 7-th " " :		$V(2,1) = V^{-1}(1,2)$.

In this way, all seven relations (III′) have been used, each to eliminate one generator from (II″), expressing all by the remaining three generators

$$Z = S^3 = \begin{pmatrix} 1 & 3 \\ 0 & 1 \end{pmatrix}; \qquad U = TS^3T^{-1} = \begin{pmatrix} 1 & 0 \\ -3 & 1 \end{pmatrix};$$

$$V(1,2) = TSTS^2TS^{-1}T^{-1}S^{-2}T^{-1} = \begin{pmatrix} 4 & 3 \\ -3 & -2 \end{pmatrix},$$

and no supplementary defining relations are left. It follows that: *the subgroup* $\Gamma_0^0(3)$ *is the free group of three generators.*

6. 3. In the general case, with $p > 3$, we proceed in a similar way:

By (43), the 1-st of (III): $W(1) = Z^{-1}$;

by (42), with $k = p-1$, $k_* = 1$: $Y(p-1) = W^{-1}(1) = Z$;

by (48), the 4-th in (III), which we may write, conveniently as

$Y(p-k) = Y(p-k+1) \cdot V^{-1}(k-1, \bar{l})$, or, changing k into $p-k$:
$Y(k) = Y(k+1)V^{-1}(p-k-1 \cdot \bar{l}')$ we eliminate,[12] successively all $Y(k)$ for $1 \leqq k \leqq p-2$.

By (42), with $1 \leqq k \leqq p-2$, we eliminate the generators $W(k_*)$, with $2 \leqq k_* \leqq p-1$.

By (45), the third in (III), written for $2 \leqq k \leqq p-1$, we eliminate the generators $X(k)$, except $X(p-1)$;

by (44), finally, the second in (III), we eliminate also $X(p-1)$.

In this way, we have used up all relations (III), except the last two, (i. e., all except (41) and (56)) and have eliminated all generators, except Z, U and the $V(k, l)$. As Z and U do not appear in (41) or (56), they will not be eliminated and enter in no defining relation.

7. 1. Our problem has been reduced to the study of the system

$$(41) \quad V(k,l) \cdot V(k_1, l_*) = I \qquad \begin{matrix} 1 \leqq k \leqq p-1 \\ 1 \leqq l \leqq p-1; kl \not\equiv 1. \end{matrix}$$

$$(56) \quad V(k,l) \cdot V(k_1, l_1) \cdot V(k_2, l_2) = I \qquad \begin{matrix} 1 \leqq k \leqq p-1; kl \not\equiv 1 \\ 2 \leqq l \leqq p-1; k(l-1) \not\equiv 1, \end{matrix}$$

formed by the "binary" relations (41) and the "ternary" relations (56). In (41), for each of the $p-1$ values which k may take, l can take $p-2$ distinct values, as $kl \equiv 1$ is excluded. But the $(p-1)(p-2)$ relations obtained in this way from (41) are not all distinct. In fact, by (36),

$$(k_1)_1 \equiv l_*(k_1 l_* - 1) \equiv l_*\{l(kl-1)l_* - 1\} \equiv l_*(-kl+1-1) \equiv k$$

and $(l_*)_* = l$, so that

$V(k_1, l_*) \cdot V((k_1)_1, (l_*)_*) = I$ is the same as $V(k_1, l_1) \cdot V(k, l) = I$.

Therefore there are, generally, only $\frac{1}{2}(p-1)(p-2)$ distinct relations (36). It may happen, however, that $V(k, l) = V(k_1, l_*)$, namely if, and only if

$$(59) \qquad l = l_* \qquad \text{and} \qquad (60) \qquad k = k_1.$$

[12] By (49) it results, that $\bar{l}' = 1 + (k+1)_*$.

(59) means that l is a solution of the congruence

$$\lambda^2 \equiv -1. \tag{61}$$

This congruence has no solution if $(-1/p) = -1$, i. e., if $p \equiv -1 \pmod 4$; or has two solutions, if $(-1/p) = +1$, i. e., if $p \equiv 1 \pmod 4$. Let us assume that $(-1/p) = +1$, so that there are two solutions, λ'_1 and λ'_2 of (61). Then, (60) becomes: $k \equiv \lambda'(k\lambda' - 1)$, or, by (61),

$$2k \equiv -\lambda'. \tag{62}$$

As $\lambda' \not\equiv 0$, (62) has, for each, λ'_1 and λ'_2, unique solutions satisfying $1 \leqq k \leqq p-1$. Therefore, if $(-1/p) = +1$, two of the binary relations (41) are of the form

$$V^2(k'_1, \lambda'_1) = I, \qquad V^2(k'_2, \lambda'_2) = I, \tag{63}$$

and $\frac{1}{2}\{(p-2)(p-1) - 2\} = \frac{1}{2}(p-2)(p-1) - 1$ distinct binary relations, with $V(k, l) \not\equiv V(k_1, l_1)$ are left. If $(-1/p) = -1$, no "cyclic" relations (63) exist and there are, as seen, $\frac{1}{2}(p-2)(p-1)$ distinct, non-cyclic, binary relations (41). In all cases, therefore, there are

$$\tfrac{1}{2}\{(p-2)(p-1) - (1 + (-1/p))\} \tag{64}$$

distinct, binary relations, different from the (63).

7. 2. Concerning the ternary relations (56), as $2 \leqq l \leqq p-1$, we have $p-2$ choices for l; on account of the two conditions, there are, for each l, $p-3$ choices for k; there exist, therefore, $(p-2)(p-3)$ ternary relations. But, we observed already [13] that the three relations

$$V(k, l) \cdot V(k_1, l_1) \cdot V(k_2, l_2) = I, \qquad V(k_1, l_1) \cdot V(k_2, l_2) \cdot V(k_3, l_3) = I,$$
$$V(k_2, l_2) \cdot V(k_3, l_3) \cdot V(k_4, l_4) = I,$$

are not distinct; therefore, if for every pair of numbers (k, l), we have [14] $(k_1, l_1) \not\equiv (k, l)$, then there are exactly $(p-2)(p-3)/3$ distinct ternary relations, among distinct generators. If, however, for some pair of values $(k, l) = (k_1, l_1)$ then also $(k, l) = (k_1, l_1) = (k_2, l_2)$ as, on account of (54), (55), k_2 and l_2 are found from (k_1, l_1) in the same way, as (k_1, l_1) are found from (k, l). It results that in (56) either all three generators are distinct, or else, all three are equal, and the corresponding relation from (56) becomes

$$V^3(k, l) = I. \tag{65}$$

---

[13] See remark at end of paragraph **5. 4. 3.**

[14] Here $(k_1, l_1) = (k, l)$ means $k_1 = k$ and $l_1 = l$; $(k_1, l_1) \neq (k, l)$ means that at least one of the letters is not the same in both pairs.

The equality $(k, l) = (k_1, l_1)$ means explicitly

(66) $l_1 = l_* + 1 = l$ *and* (67) $k_1 = l(kl - 1) = k.$

Multiplying (66) by l, we obtain, successively

(68) $-1 + l \equiv l^2$ whence $4l^2 - 4l \equiv -4$, and

(69) $(2l - 1)^2 \equiv -3.$

The congruence (69) has no solution, if $(-3/p) = -1$, i. e., if $p \equiv -1 \pmod 3$ and then, the (56) reduce to $(p-2)(p-3)/3$ distinct, non-cyclical, ternary relations, each containing three different generators. If $(-3/p) = +1$, i. e., if $p \equiv 1 \pmod 3$, then (69) has two distinct solutions; let us call them λ''_1 and λ''_2. With these values, (67) becomes $\lambda''(k\lambda'' - 1) \equiv k$, whence $k(\lambda''^2 - 1) \equiv \lambda''$, or, by (68), $k(\lambda'' - 2) \equiv \lambda''$, and multiplying by $-\lambda''_*$,

(70) $k(1 + 2\lambda''_*) \equiv 1.$

For $p > 3$, $1 + 2\lambda''_* \not\equiv 0$. In fact, from $1 + 2\lambda''_* \equiv 0$ we can conclude, successively to

$$2\lambda''(2\lambda''_* + 1) \equiv 0, \quad -4 + 2\lambda'' \equiv 0, \quad 2\lambda'' - 1 \equiv 3 \text{ and } (2\lambda'' - 1)^2 \equiv 9,$$

so that, by (69),

(71) $12 \equiv 0 \pmod p$, impossible for $p > 3$.

For $p > 3$, (70) has, therefore, exactly one solution satisfying $1 \leqq k \leqq p - 1$, for each of the two solutions λ''_1 and λ''_2 of (69). Let us call the corresponding values of k, k''_1 and k''_2 respectively. Consequently, for $p > 3$ and $(-3/p) = +1$, there are two generators, $V(k''_1, \lambda''_1)$ and $V(k''_2, \lambda''_2)$, which satisfy cyclic relations (65), of third order. The remaining $(p-2)(p-3) - 2$ relations (56) are, by a previous remark, 3 by 3 identical, so that there are, in this case $\{(p-2)(p-3) - 2\}/3$ distinct, non-cyclical, ternary relations. As, for $(-3/p) = -1$ we have $(p-2)(p-3)/3$ such relations, it results that, for every p, there are

(72) $\frac{1}{3}\{(p-2)(p-3) - (1 + (-3/p))\}$

distinct, ternary relations, different from (65).

7. 3. The total number of distinct relations, which can be used for the elimination of generators $V(k, l)$ is, by (64) and (72),

$$\tfrac{1}{2}\{(p-2)(p-1) - (1 + (-1/p))\} + \tfrac{1}{3}\{(p-2)(p-3) - (1 + (-3/p))\}.$$

Each one of these relations can be used,[15] in order to eliminate one generator. As there are $(p-1)(p-2)$ different[16] generators $V(k,l)$, we are left, after using all non-cyclical relations, with

$$N=(p-1)(p-2)-\tfrac{1}{2}\{(p-2)(p-1)-(1+(-1/p))\}$$
$$-\tfrac{1}{3}\{(p-2)(p-3)-(1+(-3/p))\}$$

independent generators. A simple transformation reduces the last expression to

$$N=(p^2+p-1)/6+(3(-1/p)+2(-3/p))/6. \tag{73}$$

Here, the second term is periodic, with period 12, and it is easily found that[17]

$$3(-1/p)+2(-3/p)=\begin{cases} 5 & \text{if } p\equiv 1 \pmod{12} \\ 1 & \text{“ } p\equiv 5 \text{ “} \\ -1 & \text{“ } p\equiv 7 \text{ “} \\ -5 & \text{“ } p\equiv 11 \text{ “} \end{cases}.$$

Herefrom it follows, by (73) that

$$6N=\begin{cases} p^2+p+4 & \text{if } p\equiv 1 \pmod{12} \\ p^2+p & \text{“ } p\equiv 5 \text{ “} \\ p^2+p-2 & \text{“ } p\equiv 7 \text{ “} \\ p^2+p-6 & \text{“ } p\equiv 11 \text{ “} \end{cases} \tag{74}$$

so that, in all cases

$$N=2[(p+2)(p-1)/12]+1 \text{ [18]} \tag{75}$$

To this number of independent generators $V(k,l)$ of $\Gamma_0^0(p)$, we still have to add the generators $Z=S^p$ and $U=TS^pT^{-1}$. Therefore, the total number of independent generators of $\Gamma_0^0(p)$ is

$$m=N+2=2[(p+2)(p-1)/12]+3. \tag{76}$$

7.4. In the process of eliminating generators, all non-cyclic relations

---

[15] We suppress here the proof of the statement, as it reduces to the elementary but lengthy exposition of a systematical method of elimination, patterned exactly after the corresponding methods in [4] and [7].

[16] There are $p-1$ possible choices for k, $1\leqq k\leqq p-1$; and for each k, there are $p-2$ different possible choices for l, satisfying: $1\leqq l\leqq p-1$ and $kl\not\equiv 1$.

[17] A prime $p>3$ can be only of one of the four forms considered here.

[18] Square brackets represent here, as usual, the largest integer, not exceeding the enclosed quantity.

have been used. As defining relations there remain, therefore, only (63) and (65), which, evidently, cannot be used to eliminate generators. Recapitulating, the defining relations,

$$V^2(k'_1, \lambda'_1) = I, \qquad V^2(k'_2, \lambda'_2) = I \tag{63}$$

exist, if $p \equiv 1 \pmod 4$, λ'_1, λ'_2, k'_1, k'_2 being the solutions of (61) and 62), respectively.

$$V^3(k''_1, \lambda''_1) = I, \qquad V^3(k''_2, \lambda''_2) = I \tag{65}$$

exist, if $p \equiv 1 \pmod 3$, λ''_1, λ''_2, k''_1, k''_2 being the solutions of (69) and (70), respectively.[19] If $p \equiv 1 \pmod{12}$, then both (63) and (65) exist; if $p \equiv -1 \pmod{12}$, then there are no defining relations at all and $\Gamma_0^0(p)$ is the free group of m generators. The number of defining relations is, therefore, for every p,

$$\nu = \{1 + (-1/p)\} + \{1 + (-3/p)\} = 2 + (-1/p) + (-3/p), \tag{77}$$

as results immediately, either from **7. 3**, or as a consequence of (64) and (72).

7. 5. The explicit form of the independent generators of $\Gamma_0^0(p)$ is easily determined, using the definitions (II), and we find

$$Z = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix}, \qquad U = \begin{pmatrix} 1 & 0 \\ -p & 1 \end{pmatrix}, \tag{78}$$

$$V(k, l) = \begin{pmatrix} k_1(1 + ll_*) - l & 1 + ll_* \\ k_1 l_* + kl - 1 - kk_1(1 + ll_*) & -k(1 + ll_*) + l_* \end{pmatrix}.$$

8. 1. As is well-known[20] $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is parabolic if, and only if $a + d = \pm 2$. We see, therefore, from (78) that Z and U are parabolic for every p. As for $V(k, l)$, it is parabolic if, and only if:

$$\begin{gathered} k_1(1 + ll_*) - l - k(1 + ll_*) + l_* = \pm 2 \text{ or} \\ (k_1 - k)(1 + ll_*) + (l_* - l) = \pm 2. \end{gathered} \tag{79}$$

As $1 + ll_* = n \cdot p$, $n = 1, 2, \cdots$ and $|l_* - l| < p$, (79) can hold only if either:

(i) $k_1 - k = 0$ and $|l_* - l| = 2$

[19] By elementary computations it is easy to show also that:

$$\lambda'_1 + \lambda'_2 = k'_1 + k'_2 = p, \qquad \lambda''_1 + \lambda''_2 = p + 1 \text{ and } k''_1 + k''_2 = p,$$

and this remark simplifies numerical computations.

[20] See, e. g., [1], page 21.

or if:

(ii) $|k_1 - k| = 1$ and $ll_* + 1 = p$, whence, by (79), $|l_* - l| = p - 2$.

The first alternatve is impossible. In fact, substituting k_1 by its value and multiplying by l_*, $k_1 - k = 0$ becomes:

$$k(l_* + l) \equiv 1; \text{ also, from } |l_* - l| = 2 \text{ results:}^{21} \tag{80}$$

$$k(l_* - l) \equiv \pm 2k. \tag{81}$$

Adding, we obtain: $2kl_* = 1 \pm 2k$, whence, after simple transformations:

$$2k(l \pm 1) \equiv \mp l. \tag{82}$$

Subtracting (81) from (80) we obtain: $2kl \equiv 1 \mp 2k$, so that:

$$2k(l \pm 1) \equiv 1. \tag{83}$$

Comparing (82) with (83), it results that $l = 1$, or $l = p - 1$, leading, respectively, to $l_* = p - 1$, or $l_* = 1$, so that $|l_* - l| = p - 2 \neq 2$, proving that the first alternative is impossible.

In the second alternative, the system $ll_* = p - 1$, $l_* - l = \pm (p - 2)$ is readily solved and has the solutions

$$l = 1, \quad l_* = p - 1 \text{ and } l = p - 1, \quad l_* = 1.$$

As both solutions actually lead to $k_1 = k - 1$, $k_1 = k + 1$, respectively, the corresponding $V(k, 1)$ and $V(k, p - 1)$ are, effectively, parabolic.

8. 2. By (41), $V(k, 1) = V^{-1}(k_1, p - 1)$ and no parabolic generators of the form $V(k, 1)$ need appear in the set of independent generators of $\Gamma_0^0(p)$. Of the $p - 2$ parabolic generators. $V(k_1, p - 1)$ $(k_1 = 1, 2, \cdots, p - 2)$, we can eliminate all but one, using the relations (56). For that, let us list them, starting with $l = 2$; to it corresponds $l_* = (p - 1)/2$ and $l_1 = l_* + 1 = (p + 1)/2$; $l_{1*} = p - 2$ and $l_2 = l_{1*} + 1 = p - 1$. Therefore, for $V(k, 2)$, taken as first factor in (56), the last factor is $V(k_2, p - 1)$, where k_2 takes all integral values $1 \leqq k \leqq p - 2$, except $(p - 1)/2$, as $(p - 1)/2 \cdot (p - 2) = p(p - 3)/2 + 1 \equiv 1$, contradicting (53). From preceding relations it is also easy to see, that in any ternary relations (56), there can be at most one parabolic factor; consequently, we can use the $p - 3$ relations, starting with $V(k, 2)$, in order to eliminate $p - 3$ of the $p - 2$ parabolic generators $V(k, p - 1)$ and are left only with $V(\frac{1}{2}(p - 1), p - 1)$, which cannot be eliminated. Consequently, it follows that among the independent generators

---

[21] Here, whenever double signs are written, we have to consider either all upper, or all lower signs.

of every subgroup $\Gamma_0^0(p)$, $p > 3$, there are exactly three parabolic generators, namely:

$$Z = S^p = \begin{pmatrix} 1 & p \\ 0 & 1 \end{pmatrix}, \qquad U = TS^pT^{-1} = \begin{pmatrix} 1 & 0 \\ -p & 1 \end{pmatrix} \text{ and}$$

$$V(\tfrac{1}{2}(p-1), p-1) = \begin{pmatrix} \frac{1}{2}p(p-1)+1 & p \\ -\frac{1}{4}p(p-1)^2 & -\frac{1}{2}p(p-1)+1 \end{pmatrix}.$$

9. Using the relations (61), (62), (63), (65), (69), (70) and (76), the following results are found for the subgroups $\Gamma_0^0(p)$, corresponding to the first few values of p:

	Generators	*Defining Relations*
$p=2$	Z, U	none
$p=3$	$Z, U, V(1,2)$	none
$p=5$	$Z, U, V(2,4); V(4,2), V(1,3)$ $V(2,2), V(4,3)$	$V^2(4,2)=I, V^2(1,3)=I.$
$p=7$	$Z, U,$ other 9 indep. gen. $V(k,l)$	$V^3(3,3)=I, V^3(4,5)=I.$
$p=11$	$Z, U,$ " 21 " " $V(k,l)$	none
$p=13$	$Z, U,$ " 31 " " $V(k,l)$	$V^2(4,5)=I, V^2(9,8)=I,$ $V^3(11,10)=I, V^3(2,4)=I.$
$p=17$	$Z, U,$ " 57 " " $V(k,l)$	$V^2(15,4)=I, V^2(2,13)=I.$
$p=19$	$Z, U,$ " 63 " " $V(k,l)$	$V^3(5,12)=I, V^3(14,8)=I.$
$p=23$	$Z, U,$ " 91 " " $V(k,l)$	none.

10. The genus of the fundamental region. Let us consider, for a given subgroup $\Gamma_0^0(p)$, a fundamental region [22] such, that the independent generators found in paragraphs 6-10 should establish the correspondence between pairs of its sides.

Let $2n$ be the number of sides; k the number of cycles, corresponding to fixed point (that is, exclusive of cycles of accidental corners); m and ν the number of independent generators and defining relations of $\Gamma_0^0(p)$ respectively; finally g the genus of the fundamental region, as defined. Then [23] $2n = 4g + 2k$, hence

$$g = \tfrac{1}{2}(n-k). \tag{84}$$

[22] See, e. g., [2], vol. 2, p. 4.

[23] See [2], vol. 2, p. 303 and vol. 1, p. 319. (The notations there are different from ours.)

According to the definition of the fundamental region,

$$n = m. \tag{85}$$

The number k, of cycles, equals the number of defining relations ν, to which we have to add the number of parabolic generators; as shown in **8.2** for every $p > 3$, there are exactly three independent parabolic generators. Therefore:

$$k = \nu + 3. \tag{86}$$

The values of m and ν are known from (76) and (77). Substituting these values in (85), (86) and then in (84), we obtain, for $p > 3$, the following expression for g:

$$\begin{aligned} g &= [(p+2)(p-1)/12] - \{1 + \tfrac{1}{2}((-1/p) + (-3/p))\} \\ &= [(p+2)(p-1)/12] - K(p). \end{aligned} \tag{87}$$

Here $K(n)$ is a periodic function, of period 12, such that $K(n) = 0, 1, 2$, according as $n \equiv -1, \pm 5, +1 \pmod{12}$, respectively. As every prime number $p > 3$ satisfies one of the four congruences $p \equiv r \pmod{12}$, $-6 \leqq r \leqq 6$, with $r = \pm 1, \pm 5$ it results that $K(p) = 1 - r(r^2 - 25)/24$; so that

$$g = [(p+2)(p-1)/12] - 1 + r(r^2 - 25)/24. \tag{87'}$$

For $p = 2$ and $p = 3$ we find directly, by (84), (85) and (86) with the known values for m and ν, that $g = 0$ in both cases. (87') is valid for $p > 3$ and using it, we find as genus of the fundamental regions, corresponding to the subgroups listed in paragraph **9**:

p:	2,	3,	5,	7,	11,	13,	17,	19,	23,	$\cdots$;
g:	0,	0,	1,	2,	10,	13,	24,	30,	45,	$\cdots$.

11. We can use the results obtained so far for subgroups $\Gamma_0^0(p)$, in order to establish some properties of the "principal subgroups"[24] $\Gamma(p)$ defined by (4), (4'). Evidently the $\Gamma(p)$ are subgroups of the $\Gamma_0^0(p)$. We may, therefore, in order to determine their structure, apply the same method used in paragraphs **3** to **8**, starting with the $\Gamma_0^0(p)$ as known groups, given by their generators and defining relations, as listed in paragraph **9**. In perfect analogy with (14), (15), (16), etc., we write now:

[24] See [7].

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \varepsilon \Gamma(p) \qquad \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = R \varepsilon \Gamma_0^0(p)$$

(88)

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} R \varepsilon \Gamma_0^0(p),$$

and have, on account of (4) and (4′):

(89) $$b \equiv c \equiv \beta \equiv \gamma \equiv B \equiv C \equiv 0 \qquad a \equiv d \equiv 1.$$

Writing explicitly (88) we obtain:

(90)
$$\begin{aligned} a\alpha + b\gamma &\equiv A \\ a\beta + b\delta &\equiv B \\ c\alpha + d\gamma &\equiv C \\ c\beta + d\delta &\equiv D. \end{aligned}$$

On account of (1) and (89), the first of (90) reduces to

(91) $$\alpha \equiv A;$$

the others reduce to trivial identities.

Let[25] $A \equiv k$, $1 \leqq k \leqq \frac{1}{2}(p-1)$. For $k = 1$, (88) is satisfied by $R_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$, and, in general, for $2 \leqq k \leqq \frac{1}{2}(p-1)$, by: $R_k = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ with $\alpha \equiv k$, $\beta \equiv \gamma \equiv 0$, $\delta \equiv p - k$. With these notations, we obtain the relation, similar to (7′):

(92) $$\Gamma_0^0(p) = \sum_{n=1}^{(p-1)/2} \Gamma(p) \cdot R_k,$$

with $R_1 = I$.

For $p = 2$ or $p = 3$, (92) becomes, simply:

$$\Gamma_0^0(2) = \Gamma(2), \qquad \Gamma_0^0(3) = \Gamma(3), \text{ respectively.}$$

For $p > 3$ and $p \equiv -1 \pmod{12}$ we have seen in **7.4** that $\Gamma_0^0(p)$ is the free group of m generators, with m determined by (76). As $\Gamma(p)$ is a subgroup of $\Gamma_0^0(p)$, it results, by a theorem of J. Nielsen [3] and O. Schreier [6] that, choosing conveniently its generators, $\Gamma(p)$ is also a free group. The number N, of the corresponding independent generators, can be found, using a relation due also to O. Schreier [6], namely (in slightly modified notations, to correspond with ours):

---

[25] As $AD \equiv 1$, the case $A \equiv 0$ can be discarded from the start. If $\frac{1}{2}(p-1) < k \leqq p-1$, then consider $k - p$ and use $\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -a & -b \\ -c & -d \end{pmatrix}$.

$$N = j(m-1)+1, \tag{93}$$

where m is the number of generators of $\Gamma_0^0(p)$, as given by (76) and j is the index of $\Gamma(p)$ in $\Gamma_0^0(p)$. By (92), $j = \frac{1}{2}(p-1)$; substituting this value and that of m in (93), taking also into account that $p \equiv -1 \pmod{12}$, we obtain:

$$N = p(p^2-1)/12+1,$$

a value found by H. Frasch without the restriction $p \equiv -1 \pmod{12}$.

UNIVERSITY OF PENNSYLVANIA.

BIBLIOGRAPHY.

[1] L. R. Ford, *Automorphic Functions*, New York, 1929.

[2] R. Fricke und F. Klein, *Vorlesungen über die Theorie der automorphen Funktionen*, Leipzig, 1897.

[3] J. Nielsen, "On Regning med ikke-kommutative Faktorer og dens Anvendelse i Gruppeteorien," *Matematisk Tidskrift*, Afd. B (1921), pp. 77-94.

[4] H. Rademacher, "Ueber die Erzeugenden von Kongruenzuntergruppen der Modulgruppe," *Abhandlungen aus dem Mathematischen Seminar der Hamburgischen Universität*, vol. 7 (1929), pp. 134-148.

[5] K. Reidemeister, "Knoten und Gruppen," *ibid.*, vol. 5 (1926), pp. 7-23.

[6] O. Schreier, "Die Untergruppen der freien Gruppen," *ibid.*, vol. 5 (1926), pp. 161-183.

[7] H. Frasch, "Die Erzeugenden der Hauptkongruenzgruppen für Primzahlstufen," *Mathematische Annalen*, vol. 108 (1933), pp. 229-252.

ON THE SOLUTION OF SUM-EQUATIONS.*

By I. M. SHEFFER.

1. Introduction. The system of linear equations

$$b_{n0}x_0 + b_{n1}x_1 + \cdots + b_{nj}x_j + \cdots = c_n, \qquad (n = 0, 1, \cdots), \tag{1.1}$$

can, in many interesting cases, be reduced to a system of *sum-equations* [1]

$$\mathcal{A}: \ a_{n0}x_n + a_{n1}x_{n+1} + \cdots + a_{nj}x_{n+j} + \cdots = c_n \quad (n = 0, 1, \cdots). \tag{1.2}$$

Letting X denote the sequence $\{x_n\}$, the left side of (1.2) will be abbreviated $A_n[X]$. It is our purpose to study the system (1.2), in which all quantities are complex numbers.[2]

Define the functions $\{A_n(t)\}$ by

$$A_n(t) = \sum_{j=0}^{\infty} a_{nj}t^j \qquad (n = 0, 1, \cdots). \tag{1.3}$$

A sequence $\{H_n(1/t)\}$ of polynomials in $1/t$ is introduced as biorthogonal to the set $\{t^nA_n(t)\}$, and this defines the generating function

$$\Phi(u, 1/t) = \sum_{n=0}^{\infty} H_n(1/t)u^n. \tag{1.4}$$

The function Φ plays a central rôle in the theory. Use is also made of the functions

$$\Psi(1/u, t) = \sum_{n=0}^{\infty} t^nA_n(t)/u^n \tag{1.5}$$

and

$$\Lambda(u, 1/t) = \sum_{n=0}^{\infty} H_n(1/t)u^nA_n(u). \tag{1.6}$$

The results of § 2 are related to work of von Koch [9]. They state conditions under which (1.2) has a unique solution, and set the background for § 3 where it is shown how from knowledge of the analytic continuation

* Received January 7, 1950; revised May 2, 1950. Presented to the American Mathematical Society, October, 1949.

[1] For example, if all principal minors in (1.1) are different from zero, then (1.2) can be obtained by forming linear combinations of the equations of (1.1) in an obvious manner.

[2] Many articles have been written on sum-equations, including all the references (save that of Boas) cited in the bibliography at the end of the present work.

of the function $\Phi(u, 1/t)$ (in the variable u) one obtains solutions of the homogeneous system (1.2) $[c_n \equiv 0]$.

2. Preliminary results. We introduce some definitions.

Definition. System $\mathcal{A}$ has the property $P_1[R]$ if $a_{n0} \neq 0$ for all n in (1.3) and if each series of (1.3) has a convergence radius at least as great as R $(0 < R \leqq \infty)$.

Definition. A series $\Sigma f_{j_1,\cdots,j_k}(u_1, \cdots, u_p)$ is said to *converge "uniformly"* for u_s on a point set $\mathfrak{F}_s$ $(s = 1, \cdots, p)$ if it converges uniformly (in the usual sense) for u_s on $\mathfrak{F}^*_s$ where $\mathfrak{F}^*_s$ is an arbitrary closed point set contained in $\mathfrak{F}_s$.

LEMMA 2.1. *Let condition $P_1[R]$ hold. For each $n = 0, 1, \cdots$ there is a unique polynomial $H_n(1/t)$, of degree exactly n in $1/t$, that is orthogonal*[3] *to the set $\{t^n A_n(t)\}$ in the sense*

$$\frac{1}{2\pi i}\int_C H_n(1/t)\, t^k A_k(t)\, dt/t = \delta_{nk}. \tag{2.1}$$

Here C is any rectifiable simple closed curve[4] *surrounding $t = 0$ and lying in $|t| < R$. Moreover,*

$$a_{0n}H_0(1/t) + a_{1,n-1}H_1(1/t) + \cdots + a_{n0}H_n(1/t) = 1/t^n. \tag{2.2}$$

The proof of (2.1) is direct, and (2.2) is a consequence of (2.1).

Remark. Let n be fixed. If $H_n(1/t) = \sum_{j=0}^{n} h_{nj}/t^j$ then we see from (2.1) and (2.6) that $\{x_j\}$ defined by $x_j = h_{nj}$, $j = 0, 1, \cdots, n$; $x_j = 0$, $j > n$ is a solution of system (1.2) with $c_j = 0$ if $j \neq n$, $c_n = 1$. It is thus seen that the results of § 2 hold at least formally, especially those results involving series (2.7). It would be possible to express these results (of § 2) in terms of the matrix

$$\mathcal{A} = \begin{bmatrix} a_{00} & a_{01} & a_{02} \cdots \\ 0 & a_{10} & a_{11} \cdots \\ \cdot & \cdot & \cdot \cdot \\ \cdot & \cdot & \cdot \cdot \end{bmatrix}; \text{ e.g., if } \mathcal{H} = \begin{bmatrix} h_{00} & \cdot & \cdot \cdot \\ h_{10} & h_{11} & \cdots \\ \cdot & \cdot & \cdot \cdot \\ \cdot & \cdot & \cdot \cdot \end{bmatrix},$$

then (2.2) expresses the relation $\mathcal{H}'\mathcal{A} = \mathcal{I}$, where $\mathcal{I}$ is the identity and $\mathcal{H}'$ is the transpose of $\mathcal{H}$. For the purpose of preparing the way for § 3, however, it is desirable to use the functions Φ, Ψ.

---

[3] Boas [1] obtained a similar orthogonality relation, with a different purpose in view.

[4] Here and elsewhere, contours are traversed in the positive direction.

Definition. System $\mathcal{A}$ has the property $P_s[R, q, r]$, $s = 2, 3, 4$ if $P_1[R]$ holds and if q, r exist so that respectively: $(s = 2)$ $0 < q$, $0 \leqq r$ and the Φ-series (1.4) converges "uniformly" for $|u| < q$, $|t| > r$; $(s = 3)$ $0 \leqq q$, $0 < r \leqq R$ and the Ψ-series (1.5) converges "uniformly" for $|u| > q$, $|t| < r$; $(s = 4)$ $0 < q \leqq R$, $0 \leqq r$ and the Λ-series (1.6) converges "uniformly" for $|u| < q$, $|t| > r$.

Remark. It is undesirable for our purpose if the Φ-series has either a zero or infinite radius of convergence for a given t. In such case we can choose a sequence $\{\lambda_n\}$ of non-zero constants so that on replacing $\{H_n(1/t)\}$ by $\{\lambda_n H(1/t)\}$ the new Φ-series has a finite, non-zero convergence radius. This alteration replaces $\{A_n(t)\}$ by $\{A_n(t)/\lambda_n\}$, as we see from (2.1); so system (1.2) is replaced by an equivalent system, obtained by multiplying the n-th equation $(n = 0, 1, \cdots)$ by $1/\lambda_n$.

LEMMA 2.2. *If $\mathcal{A}$ satisfies $P_4[R, q, r]$ then $q \leqq r$ and*

$$\Lambda(u, 1/t) = t/(t - u). \tag{2.3}$$

Proof. For $|u| < q$ series (1.6) can be expanded in a power series in u for each t in $|t| > r$; and by (2.2) this series is $\sum_0^\infty (u/t)^n$, convergent for $|u| < q$, $|t| > r$.

LEMMA 2.3. *Let $\mathcal{A}$ satisfy $P_2[R, q_2, r_2]$ and $P_3[R, q_3, r_3]$, with $q_3 < q_2$. Then $P_4[R, q, r]$ holds with $q = r_3$, $r = r_2$; and "uniformly" in $|u| < r_3$, $|t| > r_2$,*

$$\frac{1}{2\pi i}\int_\Gamma \Phi(s, 1/t)\Psi(1/s, u)\,ds/s = \sum_{n=0}^\infty H_n(1/t)u^n A_n(u) = t/(t - u), \tag{2.4}$$

where Γ is a rectifiable simple closed curve surrounding $s = 0$ and lying in the ring $q_3 < |s| < q_2$. Moreover, $r_3 \leqq r_2$.

Proof. Substituting (1.4), (1.5) into (2.4), the left side reduces to (1.6), "uniformly" convergent for $|u| < r_3$, $|t| > r_2$. Lemma 2.2 then applies.

Definition. By the *type* of a sequence $\{x_n\}$ is meant the number

$$((x_n)) = \limsup |x_n|^{1/n}. \tag{2.5}$$

A sequence $X = \{x_n\}$ and the series (or function) $X(1/t) = \sum_0^\infty x_n/t^n$ are uniquely determined each by the other, and when both appear it will be understood that they bear this relation to each other. It is readily seen that if system $\mathcal{A}$ satisfies $P_1[R]$ and $\{x_n\}$ is of type $\sigma < R$, then

$$A_n[X] \equiv \sum_{j=0}^{\infty} a_{nj} x_{n+j} = \frac{1}{2\pi i} \int_C X(1/t)\, t^{n-1} A_n(t)\, dt, \tag{2.6}$$

where C *is* a rectifiable simple closed curve surrounding $t = 0$ and lying in the ring $\sigma < |t| < R$.

THEOREM 2.1. *Let* $\mathcal{A}$ *satisfy* $P_4[R; q, r]$. *If* $\{x_n\}$ *is a solution of system* (1.2), *of type less than* q, *then*

$$X(1/t) = \sum_{n=0}^{\infty} c_n H_n(1/t), \tag{2.7}$$

the series being "uniformly" convergent in $|t| > r$.

Proof. As $((x_n)) = \sigma < q \leqq R$, series (1.6) and $X(1/u) = \sum_0^{\infty} x_n/u^n$ converge uniformly for $|u| = \lambda$ $(\sigma < \lambda < q)$ and $|t| \geqq r' > r$. Hence if C is the circle $|u| = \lambda$,

$$\frac{1}{2\pi i} \int_C \Lambda(u, 1/t) X(1/u)\, du/u = \sum_{n=0}^{\infty} H_n(1/t) \cdot \frac{1}{2\pi i} \int_C X(1/u) u^{n-1} A_n(u)\, du,$$

"uniformly" convergent for $|t| > r$. From (2.3) the left side has the value $X(1/t)$, and from (2.6) the coefficient of $H_n(1/t)$ is $A_n[X]$, which is to say, c_n.

Remark. Under the conditions of Theorem 2.1 the homogeneous system (1.2) $[c_n \equiv 0]$ has no solution of type less than q other than $x_n \equiv 0$.

THEOREM 2.2. *If* $\mathcal{A}$ *satisfies* $P_1[R]$ *and* $\{c_n\}$ *is such that series* (2.7) *converges "uniformly" in* $|t| > \delta$ $(\delta < R)$ *then* $\{x_n\}$ *is a solution of system* (1.2), *and* $((x_n)) \leqq \delta$.

Proof. Multiply (2.7) by $t^{k-1} A_k(t)/(2\pi i)$ and integrate over the circle C: $|t| = \lambda$, $\delta < \lambda < R$. Using (2.1) and (2.6) we obtain $A_n[X] = c_n$; and clearly, $((x_n)) \leqq \delta$.

Remark. By using the Cauchy integral for the coefficients $\{x_n\}$ in $X(1/t)$, it is seen from (2.7) that the solution in Theorem 2.2 is given by

$$x_n = \sum_{k=n}^{\infty} \lambda_{nk} c_k \tag{2.8}$$

where

$$\lambda_{nk} = \frac{1}{2\pi i} \int_\Gamma t^{n-1} H_k(1/t)\, dt, \tag{2.9}$$

Γ being rectifiable simple closed curve surrounding $t = 0$.

THEOREM 2.3. *Let* $\mathcal{A}$ *satisfy* $P_2[R, q, r]$, *with* $r < R$. *For every* $\{c_n\}$ *with* $((c_n)) < \min\{q, R\}$, *system* (1.2) *has the solution* $\{x_n\}$ *given by*

$$X(1/t) = \tfrac{1}{2\pi i}\int_\gamma C(1/u)\Phi(u, 1/t)\,dw/u, \tag{2.10}$$

where

$$C(1/u) = \sum_{n=0}^{\infty} c_n/u^n \tag{2.11}$$

and γ *is a rectifiable simple closed curve surrounding* $u = 0$ *and lying in the ring* $((c_n)) < |u| < \min\{q, R\}$. *Moreover,* $((x_n)) \leqq r$.

Proof. The Φ series (1.4) converges uniformly for u on γ and $|t| \geqq r' > r$, so (2.10) becomes $X(1/t) = \sum_0^\infty c_n H_n(1/t)$, "uniformly" convergent for $|t| > r$. Theorem 2.2 now applies.

Remark. If $\mathcal{A}$ satisfies $P_2[R, q, r]$ then $\Phi(u, 1/t)$ can be expanded in a power series in $1/t$:

$$\Phi(u, 1/t) = \sum_{n=0}^{\infty} u^n L_n(u)/t^n, \tag{2.12}$$

"uniformly" convergent in $|u| < q$, $|t| > r$, where the functions $L_n(u)$ are analytic in $|u| < q$. In fact,

$$u^n L_n(u) = \tfrac{1}{2\pi i}\int_C \Phi(u, 1/t)\,t^{n-1}dt \tag{2.13}$$

where C is a rectifiable simple closed curve surrounding $t = 0$ and lying in $|t| > r$. The solution $\{x_n\}$ of Theorem 2.3 can therefore be expressed in the form

$$x_n = \tfrac{1}{2\pi i}\int_\gamma C(1/u)\,u^{n-1}L_n(u)\,du. \tag{2.14}$$

3. Solutions of the homogeneous system. The results of § 2 concern system (1.2) in a range where there is no solution of the homogeneous case $[c_n \equiv 0]$ other than $x_n \equiv 0$. In the present section it will be shown that non-zero solutions of the homogeneous system arise when the function $\Phi(u, 1/t)$, on being continued analytically in the variable u, has the representation (3.1) or (3.31). Thus a relation is established between the analytic properties of the function Φ on the one hand and solutions of system (1.2) on the other.

Definition. The set of functions $\Theta_j(u)$, $j = 1, \cdots, k$ is said to be "*linearly dependent*" in $|u| < q$ if constants $a_1, \cdots, a_k$ exist, not all zero,

such that the function $\sum_{j=1}^{k} a_j \Theta_j(u)$ is analytic in a circle, center $u = 0$ and radius exceeding q. In the contrary case the functions are "linearly independent" in $|u| < q$.

Definition. Let q, q_1, q' exist, with $q < q_1 < q'$, such that $\mathcal{A}$ has the property $P_2[R, q, r]$ but not $P_2[R, q_1, r]$, and let $\Phi(u, 1/t)$ on being continued analytically have the representation

$$\Phi(u, 1/t) = \sum_{j=1}^{k} R_j(1/t)\Theta_j(u) + \Phi^*(u, 1/t) \tag{3.1}$$

where the series

$$\Phi^*(u, 1/t) = \sum_{n=0}^{\infty} H^*_n(1/t)u^n \tag{3.2}$$

is "uniformly" convergent in $|u| < q'$, $|t| > r$, and the functions $H^*_n(1/t)$ are analytic in $|t| > r$. Moreover, let the functions $R_j(1/t)$ be analytic and linearly independent in $|t| > r$, and the functions $\Theta_j(u)$ analytic in $|u| < q$ and "linearly independent" in $|u| < q_1$. Then the expansion (3.1) is said to be *canonical, relative to* $[q, q_1, q'; r]$.

A canonical form is esentially unique, as the following shows.

LEMMA 3.1. *Let* $\Phi(u, 1/t)$ *have the two canonical forms*

$$\Phi(u, 1/t) = \sum_{j=1}^{k(i)} R_j^{(i)}(1/t)\Theta_j^{(i)}(u) + \Phi^{*(i)}(u, 1/t) \qquad (i = 1, 2) \tag{3.3}$$

relative to $[q, q_1, q'; r]$. *Then* (i) $k(1) = k(2)$, *and* (ii) *the members of each of the two sets* $\{R_j^{(i)}\}$ $(i = 1, 2)$ *are linear combinations of the members of the other set.*

Proof. By subtraction we have from (3.3):

$$\Phi^*(u, 1/t) = \sum_{j=1}^{k(1)} R_j^{(1)}\Theta_j^{(1)} - \sum_{j=1}^{k(2)} R_j^{(2)}\Theta_j^{(2)}, \tag{3.4}$$

where $\Phi^* = \Phi^{*(2)} - \Phi^{*(1)}$. Let $k(1) \geqq k(2)$. Since set $\{R_j^{(1)}\}$ is linearly independent, there exist $k(1)$ values $t = t_1, \cdots, t_{k(1)}$ in $|t| > r$ for which the determinant $|R_j^{(1)}(1/t_s)|$ $(j, s = 1, \cdots, k(1))$ is not zero, so on taking $t = t_1, \cdots, t_{k(1)}$ in (3.4) we can solve for $\Theta_j^{(1)}(u)$:

$$\Theta_j^{(1)}(u) = \theta_j^{(1)}(u) + \sum_{q=1}^{k(2)} \lambda_{j,q}\Theta_q^{(2)}(u) \qquad (j = 1, \cdots, k(1)), \tag{3.5}$$

where $\theta_j^{(1)}(u)$ is analytic in $|u| < q'$.

If $k(2) < k(1)$ then the matrix (λ_{jq}) is of rank less than $k(1)$, so the system of equations

$$(3.6)\qquad \sum_{j=1}^{k(1)} a_j\lambda_{jq} = 0 \qquad (q = 1, \cdots, k(2))$$

has a solution $\{a_j\}$ not identically zero. For such $\{a_j\}$ we find from (3.5) that

$$(3.7)\qquad \sum_{j=1}^{k(1)} a_j\Theta_j^{(1)}(u) = \sum_{j=1}^{k(1)} a_j\theta_j^{(1)}(u).$$

But the right side is analytic in $|u| < q'$, so the set $\{\Theta_j^{(1)}(u)\}$ is "linearly dependent" in $|w| < q_1$, contrary to hypothesis. Thus, $k(1) = k(2), = k$.

Now substitute (3.5) into (3.4), absorbing the terms containing $\theta_j^{(1)}(u)$ into Φ^*:

$$(3.8)\qquad \Phi^*(u, 1/t) = \sum_{q=1}^{k} \{\sum_{j=1}^{k} \lambda_{jq}R_j^{(1)}(1/t) - R_q^{(2)}(1/t)\}\Theta_q^{(2)}(u).$$

If there is a q and a t_1 in $|t| > r$ for which the corresponding brace is not zero, then set $\{\Theta_q^{(2)}(u)\}$ is "linearly dependent" in $|u| < q_1$, contrary to hypothesis. Hence each brace is identically zero, thus making each $R_q^{(2)}$ a linear combination of the set $\{R_j^{(1)}\}$.

THEOREM 3.1. *Let $\Phi(u, 1/t)$ have the canonical form (3.1) relative to $[q, q_1, q'; r]$, with $r < R$. Then the homogeneous system (1.2) $[c_n \equiv 0]$ has the following k linearly independent solutions of type not exceeding r:*

$$(3.9)\qquad \{x_n^{(j)}\} = \{r_{j,n}\} \qquad (j = 1, \cdots, k; n = 0, 1, \cdots),$$

where the r_{jn} are defined by

$$(3.10)\qquad R_j(1/t) = \sum_{n=0}^{\infty} r_{jn}/t^n \qquad (j = 1, \cdots, k).$$

Proof. Let C be a circle, center $t = 0$ and radius between r and R. From (3.1),

$$(3.11)\qquad \frac{1}{2\pi i}\int_C \Phi(u, 1/t)t^{n-1}A_n(t)\,dt = \sum_{j=1}^{k} \sigma_{jn}\Theta_j(u) + \sum_{j=0}^{\infty} \lambda_{jn}u^j,$$

valid for $|u| < q$. Moreover the last series converges in $|u| < q'$. Here

$$(3.12)\qquad \sigma_{jn} = \frac{1}{2\pi i}\int_C t^{n-1}A_n(t)R_j(1/t)dt; \qquad \lambda_{jn} = \frac{1}{2\pi i}\int_C t^{n-1}A_n(t)H^*_j(1/t)dt.$$

Now for $|u| < q$, Φ also has the expansion (1.4), since $\mathcal{A}$ satisfies $P_2[R, q, r]$; so for $|u| < q$ the left side of (3.11) has the value u^n. That is,

$$(3.13)\qquad \sum_{j=1}^{k} \sigma_{jn}\Theta_j(u) = u^n - \sum_{j=0}^{\infty} \lambda_{jn}u^j \qquad (|u| < q).$$

But the right side is analytic in $|u| < q'$, and the set $\{\Theta_j\}$ is "linearly independent" in $|u| < q_1$. It follows that

$$\sigma_{jn} = 0 \qquad (j = 1, \cdots, k; n = 0, 1, \cdots) ; \tag{3.14}$$

and from (3.12) and (2.6) we conclude that each sequence $\{r_{jn}\}$ $(j = 1, \cdots, k)$ is a solution of system (1.2) $[c_n \equiv 0]$.

Remarks. (i) Even if the functions $\{R_j(1/t)\}$ are not linearly independent the sequences $\{r_{jn}\}$ are solutions of system (1.2) $[c_n \equiv 0]$.

(ii) From (3.12), (3.13), (3.14) there follows the orthogonality relation

$$\tfrac{1}{2\pi i}\int_C H^*_j(1/t)t^n A_n(t)\,dt/t = \delta_{jn}, \tag{3.15}$$

where C is any rectifiable simple closed curve surrounding $t = 0$ and lying in the ring $r < |t| < R$.

It would be of interest to know if Theorem 3.1 supplies all the solutions of the homogeneous system, of type not exceeding r, without further hypotheses. Such knowledge is lacking, but with an additional condition the desired conclusion can be asserted.

THEOREM 3.2. *Let the hypotheses of Theorem* 3.1 *hold, as well as condition* $P_3[R, q_2, r_2]$ *with* $r < r_2$, $q_2 < q'$. *Then the only solutions of* (1.2) $[c_n \equiv 0]$, *of type less than* r_2, *are linear combinations of the* k *solutions of Theorem* 3.1.

Proof. Let $X = \{x_n\}$ be a solution of (1.2) $[c_n \equiv 0]$, of type $\sigma < r_2$. Let C be the circle $|t| = \rho$, with $\sigma < \rho < r_2$. From (2.6),

$$0 = \tfrac{1}{2\pi i}\int_C X(1/t)t^{n-1}A_n(t)\,dt.$$

Multiply through by u^{-n} and sum:

$$0 = \tfrac{1}{2\pi i}\int_C X(1/t)\Psi(1/u, t)\,dt/t \qquad (|u| > q_2). \tag{3.16}$$

Since the $\Phi^*(u, 1/s)$-series (3.2) is "uniformly" convergent for $|u| < q'$, $|s| > r$, we may multiply (3.16) through by $u^{-1}\Phi^*(u, 1/s)$ and integrate over the circle Γ: $|u| = q_3 (q_2 < q_3 < q')$:

$$0 = \tfrac{1}{2\pi i}\int_C X(1/t)\{\tfrac{1}{2\pi i}\int_\Gamma \Psi(1/u, t)\Phi^*(u, 1/s)\,du/u\}dt/t, \tag{3.17}$$

valid for $|s| > r$. Now for $|s| > r$, $|t| < r_2$ and u on Γ, the two series

(1.5) and (3.2), the latter with t replaced by s, converge "uniformly," so the brace, to be denoted by $U(t, 1/s)$, has the value

$$U(t,1/s) = \sum_{n=0}^{\infty} t^n A_n(t) H^*_n(1/s), \tag{3.18}$$

"uniformly" convergent for $|t| < r_2$, $|s| > r$. Thus, (3.17) becomes

$$0 = \tfrac{1}{2\pi i}\int_C X(1/t)U(t,1/s)\,dt/t. \tag{3.19}$$

Now U can be expanded in a power series

$$U(t,1/s) = \sum_{n=0}^{\infty} I_n(1/s)t^n, \tag{3.20}$$

"uniformly" convergent in $|t| < r_2$, $|s| > r$; so from (3.19) and (3.20),

$$0 = \sum_{n=0}^{\infty} x_n I_n(1/s), \tag{3.21}$$

"uniformly" convergent for $|s| > r$. If we substitute (3.18 into

$$I_n(1/s) = \tfrac{1}{2\pi i}\int_C U(t,1/s)\,dt/t^{n+1} \tag{3.22}$$

we find that for $|s| > r$,

$$I_n(1/s) = a_{0n}H^*_0(1/s) + a_{1,n-1}H^*_1(1/s) + \cdots + a_{n0}H^*_n(1/s). \tag{3.23}$$

From (3.1), (3.2) and (1.4) we see that

$$H_n(1/t) = \sum_{j=1}^{k} \theta_{jn}R_j(1/t) + H^*_n(1/t), \tag{3.24}$$

valid for $|t| > r$, where

$$\Theta_j(u) = \sum_{n=0}^{\infty} \theta_{jn}u^n \qquad (j = 1, \cdots, k);$$

so

$$I_n(1/s) = \sum_{j=0}^{n} a_{j,n-j}H_j(1/s) - \sum_{p=0}^{n} a_{p,n-p}\{\sum_{j=1}^{k} \theta_{jp}R_j(1/s)\}, \tag{3.25}$$

for $|s| > r$. Using (2.2),

$$I_n(1/s) = 1/s^n - \sum_{j=1}^{k} \beta_{jn}R_j(1/s) \qquad (|s| > r) \tag{3.26}$$

where

$$\beta_{jn} = \sum_{p=0}^{n} a_{p,n-p}\theta_{jp}. \tag{3.27}$$

Returning to (3.21), we now have

$$(3.28)\qquad 0 = X(1/s) - \sum_{n=0}^{\infty} x_n \{ \sum_{j=1}^{k} \beta_{jn} R_j(1/s) \} \qquad (|s| > \max\{\sigma, r\}).$$

We wish to interchange the summation signs, but this requires justification. Since $\{R_j\}$ is a linearly independent set, given an arbitrary number $r' > r$, there exist k values $s = s_1, \cdots, s_k$ in $|s| > r'$ such that the determinant $|R_j(1/s_p)|$ $(j, p = 1, \cdots, k)$ has a value $\Delta \neq 0$. Let C_{jp} be the cofactor of $R_j(1/s_p)$ in this determinant. Then from (3.28),

$$\sum_{p=1}^{k} C_{qp} X(1/s_p) = \sum_{n=0}^{\infty} x_n \{ \sum_{j=1}^{k} \beta_{jn} (\sum_{p=1}^{k} C_{q,p} R_j(1/s_p)) \}.$$

Now the last sum on the right has the value $\delta_{qj}\Delta$, where δ_{qj} is the Kronecker symbol; so

$$(3.29)\qquad \sum_{n=0}^{\infty} x_n \beta_{qn} = 1/\Delta \sum_{p=1}^{k} C_{qp} X(1/s_p).$$

Thus, for each $q = 1, \cdots, k$ the series on the left of (3.29) converges, and may be multiplied by $R_q(1/s)$ and summed from $q = 1$ to k, the resulting series being valid for $|s| > \sigma$. The double sum in (3.28) may therefore be rearranged to give the relation

$$(3.30)\qquad X(1/s) = \sum_{j=1}^{k} \lambda_j R_j(1/s)$$

where $\lambda_j = \sum_{n=0}^{\infty} x_n \beta_{jn}$. This establishes the theorem.

There is an extension of Theorem 3.1 to the case of infinitely many solutions.

Definition. The infinite sequence of functions $\{\Theta_j(u)\}$, $j = 0, 1, \cdots$ is "linearly dependent" in $|w| < q$ if constants $\{a_j\}$ exist, not all zero, such that series $\sum_{0}^{\infty} a_j \Theta_j(u)$ converges "uniformly" in a circle $|u| < \rho$ with $\rho > q$.

Definition. $\Phi(u, 1/t)$ is said to have the canonical representation

$$(3.31)\qquad \Phi(u, 1/t) = \sum_{j=0}^{\infty} R_j(1/t)\Theta_j(u) + \Phi^*(u, 1/t)$$

relative to $[q, q_1, q'; r]$ if: i) all the conditions of the earlier definitions[5] (with respect to (3.1)) hold except the requirement that $\{R_j(1/t)\}$ be linearly independent; and ii) the series $\sum_{0}^{\infty} R_j(1/t)\Theta_j(u)$ is "uniformly" convergent in $|u| < q$, $|t| > r$.

[5] Of course, the new definition of "linear independence" is to be used.

THEOREM 3.3. *If* $\Phi(u, 1/t)$ *has the canonical representation* (3.31) *relative to* $[q, q_1, q'; r]$, *then every sequence* $\{r_{jn}\}$, $j = 0, 1, \cdots$ *is a solution of system* (1.2) $[c_n \equiv 0]$, *of type not exceeding* r.

Proof. As with Theorem 3.1.

There is also a parallel with Theorem 3.2.

THEOREM 3.4. *Let the hypotheses of Theorem* 3.3 *hold, and also condition* $P_3[R, q_2, r_2]$ *with* $r < r_2$, $q_2 < q'$. *There exists an infinite sequence of solutions* $\{r^*_{jn}\}$, *of type not exceeding* r, *such that on setting*

$$R^*_j(1/t) = \sum_{n=0}^{\infty} r^*_{jn}/t^n \qquad (j = 0, 1, \cdot : \cdot), \tag{3.32}$$

then every solution $\{x_n\}$ *of type less than* r_2 *is given by the representation*

$$X(1/t) = \sum_{j=0}^{\infty} x_j R^*_j(1/t), \tag{3.33}$$

"uniformly" convergent in $|t| > \max\{r, \sigma\}$ *where* $\sigma = ((x_n))$.

Proof. As with Theorem 3.2, through (3.28) with the sum $j = 1, \cdots, k$ replaced by $j = 0$ to ∞, to yield

$$0 = X(1/s) - \sum_{n=0}^{\infty} x_n \{ \sum_{j=0}^{\infty} \beta_{jn} R_j(1/s) \}, \tag{3.34}$$

"uniformly" convergent in $|s| > \max\{r, \sigma\}$. The β_{jn} are still given by (3.27). Now define $R^*_n(1/t)$ by

$$R^*_n(1/t) = \sum_{j=0}^{\infty} \beta_{jn} R_j(1/t), \qquad n = 0, 1, \cdots. \tag{3.35}$$

This series is "uniformly" convergent in $|t| > r$, so $R^*_j(1/t)$ is analytic in $|t| > r$ and $((r^*_{jn})) \leqq r$ $(j = 0, 1, \cdot \cdot \cdot)$. Also, since each $\{r_{jn}\}$ is a solution of (1.2) $[c_n \equiv 0]$, we can apply relation (2.6) to (3.35) to show that each $\{r^*_{jn}\}$ is also a solution. The theorem now follows from (3.34).

Remarks. (i) It is of interest to observe that the coefficients $\{x_n\}$ in the expansion (3.33) are also the coefficients in the power series (in $1/t$) for $X(1/t)$.

(ii) Theorems 3.1 and 3.2 also have parallels in the continuous case. Suppose, for example, that

$$\Phi(u, 1/t) = \int_a^b R(1/t, \omega) \Theta(u, w) d\alpha(\omega) + \Phi^*(u, 1/t),$$

where Φ^* is as before. Assume that $\alpha(\omega)$ is of bounded variation on $I: [a, b]$, and that R, Θ are continuous for $|t| > r$, ω on I, and $|u| < q$, ω on I, respectively, and for each ω on I, the functions R, Θ are analytic in $|t| > r$, $|u| < q$ respectively. Also suppose that Θ is "linearly independent" in $|u| < q_1$, in the sense that if $a(\omega)$ is continuous on I and $\int_a^b a(\omega)\Theta(u, \omega)d\alpha(\omega)$ is analytic in $|u| < \rho$ with $\rho > q_1$, then $a(\omega) \equiv 0$. It will then follow that for each ω on I the sequence $r_n(\omega)$ defined by

$$R(1/t, \omega) = \sum_{n=0}^{\infty} r_n(\omega)/t^n$$

is a solution of (1.2) [$c_n \equiv 0$], and is of type not exceeding r. And, as in Theorem 3.2 (with the appropriate condition P_3), if $\{x_n\}$ is a solution of type σ less than r, then $X(1/t)$ has the representation

$$X(1/t) = \sum_{n=0}^{\infty} x_n \{ \int_a^b \beta_n(\omega) R(1/t, \omega) d\alpha(\omega),$$

"uniformly" convergent in $|t| > \max\{r, \sigma\}$, where

$$\beta_n(\omega) = \sum_{p=0}^{n} a_{p,n-p}\theta_p(\omega), \qquad \Theta(u, \omega) = \sum_{p=0}^{\infty} \theta_p(\omega) u^p.$$

For the non-homogeneous system we have the following extension of Theorem 2.3.

THEOREM 3.5. *Let the hypotheses of Theorem* 3.1 *hold. For every sequence* $\{c_n\}$ *with* $((c_n)) < \min\{q', R\}$, *system* (1.2) *has the solution* $\{x_n\}$ *given by*

$$X(1/t) = \tfrac{1}{2\pi i} \int_\Gamma \Phi^*(u, 1/t) C(1/u) du/u, \tag{3.36}$$

Γ *being a rectifiable simple closed curve surrounding* $u = 0$ *and lying in the ring* $((c_n)) < |u| < q'$. *Moreover,* $((x_n)) \leqq r$.

Proof. Using (2.6),

$$A_n[X] = \tfrac{1}{2\pi i} \int_\Gamma C(1/u) V_n(u) du/u, \tag{3.37}$$

where

$$V_n(w) = \tfrac{1}{2\pi i} \int_C \Phi^*(u, 1/t) t^{n-1} A_n(t) dt \tag{3.38}$$

and C is the circle $|u| = \rho$, $r < \rho < R$. The function $V_n(u)$ is analytic in $|u| < q'$, and from (3.15) its value is u^n; so $A_n[X] = c_n$, $n = 0, 1, \cdots$. That $((x_n)) \leqq r$ is clear from (3.36).

Remarks. (i) A similar result holds if we assume the hypotheses of Theorem 3. 3 or of the second remark following Theorem 3. 4.

(ii) The solution (3. 36) also has the representation

$$X(1/t) = \sum_{n=0}^{\infty} c_n H^*_n(1/t), \tag{3. 39}$$

the series being "uniformly" convergent in $|t| > r$; and it is readily shown that conversely, if $\{c_n\}$ is such that (3. 39) converges "uniformly" in $|t| > \delta$ for some $\delta < R$, then $\{x_n\}$ given by $X(1/t) = \Sigma x_n/t^n$ is a solution of (1. 2) of type not exceeding δ. (Remark (i) also applies here.)

If $\Phi(u, 1/t)$ does not possess a canonical form (3. 1) or the later extensions thereof, it is yet possible in certain cases to extend Theorem 2. 3, as the following concluding result indicates.

THEOREM 3. 6. *Let $P_2[R, q, r]$ hold with $r < R$. Let there exist a rectifiable simple closed curve γ surrounding $u = 0$ and lying in $|u| > q$, with the following two properties*: i) *There is an open connected set $\mathfrak{S}$ cotnaining γ and the point $w = 0$, in which the functions $L_n(u)$ of* (2. 12) *are single-valued and regular*; ii) *$r_1 < R$ exists such that series* (2. 12) *converges "uniformly" for u on $\mathfrak{S}$ and $|t| > r_1$. Let q_1 be the minimum distance from γ to $u = 0$. Then for every $\{c_n\}$ for which $((c_n)) < q_1$, system* (1. 2) *has the solution $\{x_n\}$ given by* (2. 14). *Moreover, $((x_n)) \leqq r_1$.*

Proof. Choose r_2 so that $r < r_2 < R$. There is an M such that $|\Phi(u, 1/t)| \leqq M$ for u on γ and $|t| = r_2$, so for u on γ, $|u^n L_n(u)| \leqq M r_2^n$. Let $\{x_n\}$ be defined by (2. 14). Then $((x_n)) \leqq r_2$, and since r_2 can be chosen as close to r_1 as we please, therefore $((x_n)) \leqq r_1$. Also, for each $n = 0, 1, \cdots$,

$$A_n[X] = \sum_{j=0}^{\infty} a_{nj} \{\tfrac{1}{2\pi i} \int_{\gamma} C(1/u) u^{n+j-1} L_{n+j}(u)\, du\}.$$

The integral and sum are readily seen to be interchangeable, so

$$A_n[X] = \tfrac{1}{2\pi i} \int_{\gamma} C(1/u) \{\sum_{j=0}^{\infty} a_{nj} \cdot \tfrac{1}{2\pi i} \int_{|t|=r_2} \Phi(u, 1/t)\, t^{n+j-1} dt\} du/u$$

$$= \tfrac{1}{2\pi i} \int_{\gamma} C(1/u) \{\tfrac{1}{2\pi i} \int_{|t|=r_2} t^{n-1} \Phi(u, 1/t) A_n(t)\, dt\} du/u.$$

For $|u| < q$, and taking r_2 in $r < r_2 < R$, we see from (1. 4) that the brace has the value u^n. But by hypothesis the brace is single-valued and regular in $\mathfrak{S}$, so its value remains u^n in $\mathfrak{S}$, and in particular on γ. Thus, $A_n[X] = c_n$; that is, $\{x_n\}$ is a solution of system (1. 2).

THE PENNSYLVANIA STATE COLLEGE.

BIBLIOGRAPHY.

[1] R. P. Boas, Jr., "Basic sets of polynomials. II," *Duke Journal of Mathematics*, vol. 16 (1949), pp. 145-149.

[2] J. Horn, "Volterrasche Integralgleichungen und Summenleichungen," *Journal für reine und angewandte Mathematische*, vol. 140 (1911), pp. 120-174.

[3] ———, "Analytische Lösungen von Summengleichungen," *Archiv der Mathematik und Phys.*, ser. 3, vol. 26 (1918), pp. 132-145.

[4] O. Perron, "Zur Theorie der Summengleichungen," *Mathematische Zeitschrift*, vol. 8 (1920), pp. 159-170.

[5] ———, "Über Summengleichungen und Poincarésche Differengleichungen," *Mathematische Annalen*, vol. 84 (1921), pp. 1-15.

[6] I. M. Sheffer, "An extension of a Perron system of linear equations in infinitely many unknowns," *American Journal of Mathematics*, vol. 67 (1945), pp. 123-140.

[7] ———, "On k-periodic systems of linear equations," *Transactions of the American Mathematical Society*, vol. 63 (1948), pp. 244-313.

[8] ———, "On the theory of sum-equations," *Bulletin of the American Mathematical Society*, vol. 55 (1949), pp. 777-788.

[9] H. von Koch, "On regular and irregular solutions of some infinite systems of linear equations," *Proceedings of the Fifth International Congress of Mathematicians*, Cambridge (1912), vol. 1, pp. 352-365.

[10] J. L. Walsh, "On the solution of linear equations in infinitely many variables by successive approximations," *American Journal of Mathematics*, vol. 42 (1920), pp. 91-96.

THE GAPS IN THE ESSENTIAL SPECTRA OF WAVE EQUATIONS.*

By PHILIP HARTMAN and CALVIN R. PUTNAM.

1. Let $q = q(t)$, where $0 \leqq t < \infty$, be a continuous, real-valued function for which the differential equation

$$x'' + (\lambda + q)x = 0 \tag{1}$$

is of the Grenzpunkt type; [9], p. 238. A fixed homogeneous boundary condition at $t = 0$,

$$x(0) \cos \alpha + x'(0) \sin \alpha = 0, \tag{2}$$

determines a self-adjoint boundary value problem with a (real) spectrum $S = S(\alpha)$. The set S' of cluster points of $S(\alpha)$ is independent of α ([9], p. 251) and will be called the essential spectrum of (1).

Since S' is a closed set, its complement is open, possibly empty. Let $\Sigma(\lambda_k, \lambda^k)$ denote the canonical decomposition of the complement of S' into disjoint open intervals. Two problems will be treated in this paper. The first concerns the size of the interval (λ_k, λ^k) as $\lambda_k \to \infty$. The second concerns asymptotic estimates, as $t \to \infty$, for eigenfunctions $x = x(t, \lambda)$ belonging to eigenvalues λ of $S(\alpha)$ satisfying $\lambda_k < \lambda < \lambda^k$ for a large value of λ_k. These problems were suggested to us by Professor Wintner.

The nature of the problems is most easily understood in the case that $q(t)$ is periodic. For this case, let $L(\lambda) \geqq 0$ denote the real part of one of the characteristic exponents belonging to λ. Then $L(\lambda) = 0$ except on a half-line $(-\infty, \lambda^0)$ and a sequence of intervals $(\lambda_1, \lambda^1), (\lambda_2, \lambda^2), \cdots$, composed of the λ-values for which (1) is of unstable type. The questions treated concern upper estimates for the length of the interval (λ_k, λ^k), as $\lambda_k \to \infty$, and upper estimates for $L(\lambda)$, as $\lambda \to \infty$. The results below concerning $L(\lambda)$, which imply that $L(\lambda) \to 0$ as $\lambda \to \infty$, show that, as λ increases, any tendency for (1) to be unstable is dampened; in the sense, that if λ is large and of unstable type, it requires long t-intervals in order that the effect of the factor e^{Lt} be appreciable.

In the case that $q(t)$ is even and periodic, the behavior of $\lambda^k - \lambda_k$ can be read off from known results on the asymptotic behavior of eigenvalues of Sturm-Liouville boundary value problems; cf. Section 7, below. This pro-

* Received March 23, 1950.

cedure can be adapted also for the case that $q(t)$ is merely periodic, not necessarily even, if it is noted that the k-th interval $[\lambda_k, \lambda^k]$ is composed of those λ-values corresponding to which (1) has a solution $x = x(t)$ which vanishes exactly k times for $0 \leqq t < p$, satisfies (2) and the same boundary condition $x(p)\cos\alpha + x'(p)\sin\alpha = 0$ at $t = p$ for some $0 \leqq \alpha < \pi$, when p is the period of $q(t)$.

Of course, the results of this paper are not restricted to the cases where $q(t)$ is periodic.

2. The gaps (λ_k, λ^k) in the essential spectrum. A sufficient condition for (1) to be of the Grenzpunkt type is that $\beta < \infty$, where $\beta = \limsup q(t)$, as $t \to \infty$ ([9], p. 238). This sufficient condition does not involve $\alpha = \liminf q(t) \geqq -\infty$, as $t \to \infty$. If $q(t)$ tends to a finite limit as $t \to \infty$ (that is, $-\infty < \alpha = \beta < \infty$), then the essential spectrum S' of (1) contains the half-line $\beta \leqq \lambda < \infty$; [3], p. 71. More generally, if $q(t)$ is bounded (that is, $-\infty < \alpha \leqq \beta < \infty$), then, for every $\lambda_0 \geqq -\beta$, the interval $\lambda_0 \leqq \lambda \leqq \lambda_0 + \beta - \alpha$ contains at least one point of S'; [6], p. 613. The object of the first part of this paper will be the refinement of this last statement in the following direction:

(I) *Let $q(t)$, where $0 \leqq t < \infty$, be a real-valued, continuous, bounded function; for $\delta > 0$, let*

$$\omega(t, \delta) = \text{fin sup}\,|\,q(s) - q(t)\,| \quad \text{for} \quad |\,s - t\,| \leqq \delta; \tag{3}$$

and

$$\omega(\delta) = \limsup T^{-1} \int_0^T \omega(t, \delta)\,dt \quad \text{as} \quad T \to \infty. \tag{4}$$

Then there exists a constant C (independent of λ_0) such that, for every sufficiently large λ_0, the interval $\lambda_0 \leqq \lambda \leqq \lambda_0 + C[\omega(\pi/\lambda_0^{\frac{1}{2}}) + 1/\lambda_0]$ contains at least one point of the essential spectrum S' of (1).

This theorem has the following interpretation and consequences: Since q is bounded, the complement of S' consists of a half-line $(-\infty, \lambda^0)$ and a (possibly empty) sequence of disjoint open intervals $(\lambda_1, \lambda^1), (\lambda_2, \lambda^2), \cdots$. The theorem (I) states that the length $\lambda^k - \lambda_k$ of the k-th gap is majorized by $C[\omega(\pi/\lambda_k^{\frac{1}{2}}) + 1/\lambda_k]$ for large λ_k. In particular, if $q(t)$ is uniformly continuous for $0 \leqq t < \infty$, so that $\omega(\delta) \to 0$ as $\delta \to +0$, then the length of the gap tends to 0 as its endpoints tend to ∞.

Similarly, if $q(t)$ satisfies a uniform Lipschitz condition, $|q(s)-q(t)| \leqq \text{const.}\,|s-t|$ for $0 \leqq s < t < \infty$, so that $\omega(\delta) \leqq \text{const.}\,\delta$, then

$$\lambda^k - \lambda_k = O(1/\lambda_k^{\frac{1}{2}}). \tag{5}$$

This same majorant for $\lambda^k - \lambda_k$ is obtained if $q(t)$ satisfies the more general conditions that $q(t)$ is of bounded variation on every finite interval and that its total variation $V(T)$ on the interval $0 \leqq t \leqq T$ satisfies

$$V(T) = \int_0^T |dq(t)| = O(T) \text{ as } T \to \infty. \tag{6}$$

In order to see this, note that $\omega(t,\delta) \leqq V(t+\delta) - V(t-\delta)$ for $t \geqq \delta$, by (3). Hence, if $T \geqq \delta$,

$$\int_\delta^T \omega(t,\delta)\,dt \leqq \int_{2\delta}^{T+\delta} V(t)\,dt - \int_0^{T-\delta} V(t)\,dt \leqq \int_{T-\delta}^{T+\delta} V(t)\,dt.$$

Then, since $V(t) \leqq \text{Const.}\,t$, the right side of the last inequality is majorized by Const. δT, as $T \to \infty$, for $\delta > 0$ fixed. Consequently, by (4), it is seen that $\omega(\delta) \leqq \text{const.}\,\delta$ and so the majorant $C/\lambda_k^{\frac{1}{2}}$ of $\lambda^k - \lambda_k$ for large λ_k follows from (I).

In the few explicit cases where q is bounded and S' is known, either there is only a finite number of gaps, or $\lambda^k - \lambda_k \to 0$ and $\lambda_k \to \infty$, as $k \to \infty$. Whether or not this is true in general will remain an open question. On the other hand, if $\psi(\lambda)$ is an arbitrary decreasing function satisfying $\psi(\lambda) \to 0$ as $\lambda \to \infty$, then there exist bounded functions q, corresponding to which the complement of the essential spectrum S' of (1) consists of a half-line $(-\infty, \lambda^0)$, and a sequence of open intervals $(\lambda_1, \lambda^1), (\lambda_2, \lambda^2), \cdots$ such that $\lambda_k \to \infty$, as $k \to \infty$, but $\lambda^k - \lambda_k = O(\psi(\lambda_k))$ fails to hold; see Section 7 below. Also, examples can be given to show that if $q(t)$ satisfies (6), then (5) is the best possible estimate; in the sense that, in general, O cannot be replaced by o in (5).

If additional local smoothness conditions are imposed on q, the assertions of Theorem (I) can be refined.

(I′) *Let $q(t)$ satisfy the conditions of* (I) *and, in addition, possess a bounded derivative $q'(t)$. If $\omega^*(t,\delta)$ and $\omega^*(\delta)$ have the same significance for q' as* (3) *and* (4) *do for q, then there exists a constant C (independent of λ_0) such that, for every sufficiently large λ_0, the interval $\lambda_0 \leqq \lambda \leqq \lambda_0 + C\lambda_0^{-\frac{1}{2}}[\omega^*(\pi/\lambda_0^{\frac{1}{2}}) + 1/\lambda_0]$ contains at least one point of the essential spectrum S' of* (1).

(I″) *Furthermore, if* $q(t)$ *possesses a uniformly bounded first and second derivative and if* $\omega^{**}(t,\delta)$, $\omega^{**}(\delta)$ *belong to* q'' *as* (3), (4) *do to* q, *then the assertion of* (I′) *applies to intervals* $\lambda_0 \leqq \lambda \leqq \lambda_0 + C\lambda_0^{-1}[\omega^{**}(\pi/\lambda_0^{\frac{1}{2}}) + 1/\lambda_0^{\frac{1}{2}}]$, *for some constant* C *and all sufficiently large* λ_0.

Thus, it follows from (I′) that if $q'(t)$ satisfies a uniform Lipschitz condition, or more generally, if

$$\int_0^T |\, dq'(t)\,| = O(T),$$

as $T \to \infty$, then the length of the k-th gap (λ_k, λ^k) is majorized by $O(1/\lambda_k)$ as $\lambda_k \to \infty$. In addition, the existence of a bounded second derivative satisfying

$$\int_0^T |\, dq''(t)\,| = O(T),$$

as $T \to \infty$, implies that $\lambda^k - \lambda_k = O(1/\lambda_k^{3/2})$ as $\lambda_k \to \infty$.

3. The characteristic exponents. If $q(t)$ denotes a continuous, bounded, real-valued function, then it is known, [11], p. 391, that there exists a positive number $k = k(\lambda)$, depending on λ alone, such that

$$x^2 + x'^2 = O(e^{kt}), \tag{7}$$

as $t \to \infty$, holds for any solution $x = x(t,\lambda)$ of the differential equation (1). Let the number $K = K(\lambda)$ be defined as the greatest lower bound of the numbers k for which (7) holds (for *every* solution x of (1)). As shown in [7], p. 137, in partial answer to a question raised by Wintner [14], p. 604, if $|\, q(t)\,|$ possesses the bound M and if λ does not belong to the set S' and satisfies $|\,\lambda\,| > M$, then there exist a constant $l = l(\lambda) > 0$ and two solutions $x = x_1(t,\lambda) \not\equiv 0$ and $x = x_2(t,\lambda)$ of (1) satisfying

$$x_1^2 + x_1'^2 \leqq e^{-lt}, \qquad x_2^2 + x_2'^2 \geqq e^{lt}. \tag{8}$$

If $L = L(\lambda)$ is defined to be the least upper bound of the numbers l for which (8) holds, then it is clear from the definition of $K(\lambda)$ that $L(\lambda) \leqq K(\lambda)$, and hence any majorant of $K(\lambda)$ will also be a majorant of $L(\lambda)$. Since (1) is of the Grenzpunkt type, the solution $x_1(t,\lambda)$ is uniquely determined to within a constant factor. Thus, it is clear that $L(\lambda)$ depends on λ alone and not on the choice of the particular pair of solutions x_1 and x_2 of (1). The number $L(\lambda)$ can be called a "characteristic exponent" associated with

λ, in analogy with the terminology used in the case that $q(t)$ is periodic. It will be the purpose of the second part of the paper to obtain, for large λ, estimates of $K(\lambda)$, and hence of $L(\lambda)$, when $q(t)$ is subject to certain smoothness conditions.

In Sections **8** and **9**, the following theorems will be proved:

(II) *If $q(t)$ satisfies the conditions of* (I), *then*

$$K(\lambda) = O(\lambda^{-\frac{1}{2}}[\omega(\pi/\lambda^{\frac{1}{2}}) + 1/\lambda]) \text{ as } \lambda \to \infty. \tag{9}$$

As a consequence of (II) it is seen that, since $q(t)$ is bounded, the estimate $K(\lambda) = O(\lambda^{-\frac{1}{2}})$ is obtained. If, in addition, $q(t)$ is uniformly continuous for $0 \leqq t < \infty$, the last estimate can be sharpened to $K(\lambda) = o(\lambda^{-\frac{1}{2}})$. Also, if $q(t)$ satisfies (6), then $K(\lambda) = O(\lambda^{-1})$. Further smoothness conditions imply better estimates for $K(\lambda)$; in this direction the following will be proved:

(II′) *If $q(t)$ satisfies the conditions of* (I′), *then*

$$K(\lambda) = O(\lambda^{-1}[\omega^*(\pi/\lambda^{\frac{1}{2}}) + 1/\lambda]) \text{ as } \lambda \to \infty. \tag{10}$$

(II″) *If $q(t)$ satisfies the conditions of* (I″), *then*

$$K(\lambda) = O(\lambda^{-3/2}[\omega^{**}(\pi/\lambda^{\frac{1}{2}}) + 1/\lambda^{\frac{1}{2}}]) \text{ as } \lambda \to \infty. \tag{11}$$

4. Proof of (I). Let $x = x(t, \lambda) \not\equiv 0$ be a solution of (1) and let $N(T, \lambda)$ denote the number of zeros of x for $0 < t < T$. According to [4], a λ-interval $(\lambda, \lambda + \Delta)$ contains a point of S' if and only if

$$N(T, \lambda + \Delta) - N(T, \lambda) \to \infty \text{ as } T \to \infty. \tag{12}$$

Since it is assumed that q is bounded, there exists a constant M such that

$$|q| \leqq M, \qquad (0 \leqq t < \infty). \tag{13}$$

Let λ be chosen so large that

$$\lambda - M > \lambda/4 > 0. \tag{14}$$

Then the zeros of x and x' separate each other. Hence, if $\theta = \theta(t, \lambda)$ is a continuous determination of

$$\theta(t, \lambda) = -\arctan(x'/\lambda^{\frac{1}{2}}x), \tag{15}$$

it follows that $\pi N(T, \lambda) = \theta(T, \lambda) + O(1)$, as $T \to \infty$. Consequently, (12) holds if and only if

$$\theta(T, \lambda + \Delta) - \theta(T, \lambda) \to \infty \text{ as } T \to \infty. \tag{16}$$

It follows from (1) and from (15) that θ is differentiable and that $\theta' = \lambda^{\frac{1}{2}} + \lambda^{-\frac{1}{2}} q\{\lambda x^2/(\lambda x^2 + x'^2)\}$; so that

$$\theta(T, \lambda) = \lambda^{\frac{1}{2}} T + \lambda^{-\frac{1}{2}} \int_0^T q\{\lambda x^2/(\lambda x^2 + x'^2)\} dt + \theta(0, \lambda). \tag{17}$$

Let $(0 \leqq) t_1 < t_2 < \cdots$ be the zeros of x. Then

$$t_{k+1} - t_k \leqq 2\pi/\lambda^{\frac{1}{2}}, \tag{18}$$

by virtue of Sturm's comparison theorem and (14). The factor $\{\cdots\}$ of q in the integral in (17) will be estimated on the interval (t_k, t_{k+1}). Because of the homogeneity of this factor, x can be replaced by any solution y of (1) which vanishes for $t = t_k$.

Let ν_k, where $k = 1, 2, \cdots$, be defined by

$$t_{k+1} - t_k = \pi/\nu_k^{\frac{1}{2}}. \tag{19}$$

Clearly, $\nu_k = \lambda + q(\xi_k)$, where ν_k is some t-value satisfying $t_k \leqq \xi_k \leqq t_{k+1}$. Hence,

$$|\lambda - \nu_k| \leqq M. \tag{20}$$

A solution y of (1) which vanishes at $t = t_k$ satisfies the integral equation

$$y(t) = \sin \nu_k^{\frac{1}{2}}(t - t_k) - \nu_k^{-\frac{1}{2}} \int_{t_k}^t Q_k(s) y(s) \sin \nu_k^{\frac{1}{2}}(t - s)\, ds, \tag{21}$$

where

$$Q_k(s) = q(s) - q(\xi_k). \tag{22}$$

Let $m = m_k = \max |y(t)|$ for $t_k \leqq t \leqq t_{k+1}$; then, by (13) and (19), it follows that $m \leqq 1 + \nu_k^{-\frac{1}{2}}(\pi \nu_k^{-\frac{1}{2}}) 2Mm$. Consequently, $m \leqq 1/(1 - 2\pi M \nu_k^{-1}) \leqq 2$ if ν_k (that is, λ) is sufficiently large. Thus (21) shows that

$$y = \sin \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-1}) \tag{23}$$

and

$$y' = \nu_k^{\frac{1}{2}} \cos \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-\frac{1}{2}}) \tag{24}$$

uniformly in k, as $\lambda \to \infty$. Since (20) implies that $|\nu_k^{\frac{1}{2}} - \lambda^{\frac{1}{2}}| = O(\lambda^{-\frac{1}{2}})$, it follows from (24) that

$$y' = \lambda^{\frac{1}{2}} \cos \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-\frac{1}{2}}). \tag{25}$$

Hence, for $t_k \leqq t \leqq t_{k+1}$,

$$\lambda y^2/(\lambda y^2 + y'^2) = \sin^2 \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-1}), \tag{26}$$

as $\lambda \to \infty$. Consequently, (17) shows that, for every fixed sufficiently large λ,

$$(27)\quad \theta(T,\lambda) = \lambda^{\frac{1}{2}}T + \tfrac{1}{2}\lambda^{-\frac{1}{2}}\int_0^T q(s)\,ds - \tfrac{1}{2}\lambda^{-\frac{1}{2}}\sum_{k=1}^{n}\int_{t_k}^{t_{k+1}} q(s)\cos 2\nu_k^{\frac{1}{2}}(s-t_k)\,ds + O(\lambda^{-3/2}T),$$

as $T \to \infty$, where $n = n(T)$ is determined by $t_n \leqq T < t_{n+1}$. In (27) and similar relations below, the constant implicit in the O-term can be chosen independent of λ.

Since the definition (19) of ν_k implies

$$\int_{t_k}^{t_{k+1}} \cos 2\nu_k^{\frac{1}{2}}(s-t_k)\,ds = 0 \text{ for } k = 1, 2, \cdots,$$

it follows that in the integral over the interval (t_k, t_{k+1}) in (27), the function $q(s)$ can be replaced by $q(s) - c_k$, where $c_1, c_2, \cdots$ are arbitrary constants. If $c_k = q(\frac{1}{2}t_k + \frac{1}{2}t_{k+1})$, then

$$|\, q(s) - c_k \,| \leqq \omega(s, \pi/\lambda^{\frac{1}{2}}), \text{ for } t_k \leqq s \leqq t_{k+1},$$

by (3) and (18). Hence,

$$\theta(T,\lambda) = \lambda^{\frac{1}{2}}T + \tfrac{1}{2}\lambda^{-\frac{1}{2}}\int_0^T q(s)\,ds + O(\lambda^{-\frac{1}{2}})\int_0^T \omega(s, \pi/\lambda^{\frac{1}{2}})\,ds + O(\lambda^{-3/2}T).$$

Then, since $(\lambda+\Delta)^{\frac{1}{2}} - \lambda^{\frac{1}{2}} > \Delta\lambda^{-\frac{1}{2}}/4$ for any fixed Δ and for sufficiently large λ, and since the first integral of the last formula line is $O(T)$ as $T \to \infty$, it is clear that

$$\theta(T,\lambda+\Delta) - \theta(T,\lambda) \geqq \Delta(\lambda^{-\frac{1}{2}}T/4 + O(\lambda^{-3/2}T) + O(\lambda^{-\frac{1}{2}})\int_0^T \omega(s, \pi/\lambda^{\frac{1}{2}})\,ds + O(\lambda^{-3/2}T),$$

as $T \to \infty$; so that there exists a constant C with the property that if $\Delta \geqq C\omega(\pi/\lambda^{\frac{1}{2}}) + C/\lambda$, then (16) holds for sufficiently large λ. This proves (I).

Incidentally, since the factor 1/4 on the right of (14) can be replaced by $1-\epsilon$, where $\epsilon > 0$ is arbitrary, it follows that $\omega(\pi/\lambda^{\frac{1}{2}})$ can be replaced by $\omega(\pi/2(1-\epsilon)^{\frac{1}{2}}\lambda^{\frac{1}{2}})$ in the statement of the theorem (I).

5. Proof of (I'). The proof depends on a refinement of (26). The existence of a bounded derivative q' and (22) show that $Q_k(s) = O(\lambda^{-\frac{1}{2}})$,

as $\lambda \to \infty$, uniformly for $t_k \leqq s \leqq t_{k+1}$. Hence, the remainder terms in (23) and (24) are $O(\lambda^{-3/2})$ and $O(\lambda^{-1})$, respectively. Thus

$$y'^2 = \nu_k \cos^2 \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-\frac{1}{2}}),$$

or, since $\nu_k = \lambda + q(\xi_k) = \lambda + q(t) + O(\lambda^{-\frac{1}{2}})$,

$$y'^2\lambda^{-1} = \cos^2 \nu_k^{\frac{1}{2}}(t - t_k) + q(t)\lambda^{-1} \cos^2 \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-3/2}).$$

Hence,

$$y^2 + y'^2\lambda^{-1} = 1 + q(t)\lambda^{-1} \cos^2 \nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-3/2}).$$

Finally, the required refinement of (26) is

$$\lambda y^2/(\lambda y^2 + y'^2) = \sin^2 \nu_k^{\frac{1}{2}}(t - t_k) - q(t)\,(4\lambda)^{-1} \sin^2 2\nu_k^{\frac{1}{2}}(t - t_k) + O(\lambda^{-3/2}), \tag{28}$$

as $\lambda \to \infty$, uniformly for $t_k \leqq t \leqq t_{k+1}$.

It follows from (17) that, for fixed large λ, the function $\theta(T, \lambda)$ is the sum of the three expressions

$$\lambda^{\frac{1}{2}}T + \tfrac{1}{2}\lambda^{-\frac{1}{2}} \int_0^T q(t)\,dt - (1/8\lambda^{3/2}) \int_0^T q^2(t)\,dt, \tag{29}$$

$$-\tfrac{1}{2}\lambda^{-\frac{1}{2}} \sum_{k=1}^{n} \int_{t_k}^{t_{k+1}} q(t) \cos 2\nu_k^{\frac{1}{2}}(t - t_k)\,dt, \tag{30}$$

$$+ (1/8\lambda^{3/2}) \sum_{k=1}^{n} \int_{t_k}^{t_{k+1}} q^2(t) \cos 4\nu_k^{\frac{1}{2}}(t - t_k)\,dt + O(\lambda^{-2}T), \tag{31}$$

as $T \to \infty$, where $n = n(T)$ satisfies $t_n \leqq T < t_{n+1}$. An integration by parts and (19) show that the k-th term of the sum in (30) is

$$-\tfrac{1}{2}\nu_k^{-\frac{1}{2}} \int_{t_k}^{t_{k+1}} q'(t) \sin 2\nu_k^{\frac{1}{2}}(t - t_k)\,dt; \tag{32}$$

since $q'(t)$ can be replaced by $q'(t) - c$, where c is an arbitrary constant, it follows from the choice $c = q'(\frac{1}{2}t_k + \frac{1}{2}t_{k+1})$, that the expression (32) is

$$O(\lambda^{-\frac{1}{2}}) \int_{t_k}^{t_{k+1}} \omega^*(t, \pi/\lambda^{\frac{1}{2}})\,dt.$$

Hence, (30) is

$$O(\lambda^{-1}) \int_0^T \omega^*(t, \pi/\lambda^{\frac{1}{2}})\,dt + O(\lambda^{-1}).$$

An integration by parts for each integral in (31) shows that (31) is $O(\lambda^{-2}T)$. Arguments similar to those used at the end of the proof of (I) complete the proof of (I').

6. Proof of (I''). Another integration by parts shows that (32) is

$$(1/4\nu_k)\{q'(t_{k+1}) - q'(t_k) - \int_{t_k}^{t_{k+1}} q'' \cos 2\nu_k^{\frac{1}{2}}(t - t_k)\,dt\}.$$

Hence (30) is the sum of

$$(33) \quad -\tfrac{1}{2}\lambda^{-\frac{1}{2}}\{(1/4\lambda)\sum_{k=1}^{n}(q'(t_{k+1}) - q'(t_k)) - \sum_{k=1}^{n}(q'(t_{k+1}) - q'(t_k))(1/4\lambda - 1/4\nu_k)\}$$

and

$$(34) \quad O(\lambda^{-3/2})\sum_{k=1}^{n}\left|\int_{t_k}^{t_{k+1}} q'' \cos 2\nu_k^{\frac{1}{2}}(t - t_k)\,dt\right| = O(\lambda^{-3/2})\int_0^T \omega^{**}(t, \pi/\lambda^{\frac{1}{2}})\,dt.$$

The first term of (33) is $O(\lambda^{-3/2})|q'(t_{n+2}) - q'(t_1)| = O(\lambda^{-3/2})$, as $T \to \infty$, since $q'(t)$ is bounded. The second is $O(n\lambda^{-5/2})$ since $|q'(t_{k+1}) - q'(t_k)| = O(1/\lambda^{\frac{1}{2}})$ and $|\nu_k - \lambda| = O(1)$. Also, $n\lambda^{-\frac{1}{2}} = O(T)$, by (13), as $T \to \infty$, for a fixed large λ.

Hence, $\theta(T, \lambda)$ is the sum of (29), (34) and terms of the type $O(\lambda^{-2}T)$, as $T \to \infty$. The proof of (I'') can now be completed as above.

7. Examples. The examples mentioned in Section **2** can be obtained by the consideration of periodic functions $q(t)$. If $q(t)$ is periodic, then the essential spectrum S' of (1) is the closure of the set of λ-values corresponding to which all solutions of (1) are bounded ([12]; cf. also [8]). Hence, if $\lambda = \lambda_k$ or λ^k, then (1) possesses a periodic or half-periodic solution.

If $q(t)$ is even and of period 2π, it is easy to see that the numbers $\lambda_1, \lambda_2, \cdots$ and $\lambda^0, \lambda^1, \cdots$ can be characterized as follows: Let $\alpha_0, \alpha_1, \cdots$ and $\beta_0, \beta_1, \cdots$ denote the eigenvalues belonging to the Sturm-Liouville problem of (1) and the boundary conditions $x(0) = x(\pi) = 0$ and $x'(0) = x'(\pi) = 0$, respectively. Then $\beta_0 = \lambda^0$ and the numbers λ_k, λ^k for $k \geqq 1$ are, up to a permutation, the numbers α_{k-1}, β_k. Hence

$$\lambda^k - \lambda_k = |\alpha_{k-1} - \beta_k| \tag{35}$$

for $k = 1, 2, \cdots$.

On the other hand, if $q(t)$ has the Fourier series

$$q(t) \sim \sum_{n=1}^{\infty} a_{2n} \cos 2nt, \qquad (a_0 = 0), \tag{36}$$

then $\alpha_{k-1} = k^2 + \frac{1}{2}a_{2k} + o(1/k)$ and $\beta_k = k^2 - \frac{1}{2}a_{2k} + o(1/k)$, as $k \to \infty$; cf. [10], p. 322; also, [1], p. 11. Consequently,

$$\lambda^k - \lambda_k = |a_{2k}| + o(1/k) \text{ and } \lambda_k = k^2 + o(1), \tag{37}$$

as $k \to \infty$.

If $\psi(\lambda)$ is an arbitrary positive function which tends monotonously to 0, as $\lambda \to \infty$, then there exist continuous functions $q(t)$ for which $a_{2k} = O(\psi(k^2 - 1))$ fails to hold as $k \to \infty$. For such functions q, it follows from (31) that $\lambda^k - \lambda_k = O(\psi(\lambda_k))$ fails to hold as $k \to \infty$.

In order to obtain the second type of example mentioned in Section **2**, it is sufficient to remark that there exist continuous functions q, of bounded variation over a period, for which $a_{2k} = O(1/k)$ but $a_{2k} \neq o(1/k)$, as $k \to \infty$; [2]. For such functions q, it follows from (37) that

$$\lambda^k - \lambda_k = O(1/\lambda_k^{\frac{1}{2}}) \text{ but } \lambda^k - \lambda_k \neq o(1/\lambda_k^{\frac{1}{2}})$$

as $k \to \infty$. This shows that, if (6) is assumed, the O in (5) cannot in general be replaced by o.

8. Proof of (II). Corresponding to any solution $x = x(t) = x(t, \lambda) \not\equiv 0$ of (1), let $h(t)$ be defined by

$$h(t) = \lambda x^2(t) + x'^2(t). \tag{38}$$

It follows from (1) that $h'(t) = -2qxx'$, so that

$$h'(t)/h(t) = -2q\{xx'/(\lambda x^2 + x'^2)\}.$$

Consequently,

$$\lambda x^2 + x'^2 = \text{const. } e^{At}, \tag{39}$$

where

$$A = A(t, \lambda) = -t^{-1} \int_0^t q\{2xx'/(\lambda x^2 + x'^2)\} ds. \tag{40}$$

Let $(0 \leqq) t_1 < t_2 < \cdots$ be the successive zeros of x. Since the factor $\{\cdots\}$ of q in the integral in (40) is homogeneous in x, x', the solution x can be replaced by any solution y of (1) vanishing at $t = t_k$.

From (23) and (25), it is seen that

$$2yy'/(\lambda y^2 + y'^2) = \lambda^{-\frac{1}{2}} \sin 2\nu_k^{\frac{1}{2}}(s - t_k) + O(\lambda^{-3/2}), \tag{41}$$

as $\lambda \to \infty$, uniformly for $t_k \leqq s \leqq t_{k+1}$. Hence, for every fixed large λ,

$$A(t, \lambda) = -(1/t\lambda^{\frac{1}{2}}) \sum_{k=1}^{n} \int_{t_k}^{t_{k+1}} q(s) \sin 2\nu_k^{\frac{1}{2}}(s - t_k)\, ds + O(\lambda^{-3/2})$$

as $t \to \infty$, where $n = n(t)$ is determined by $t_n \leqq t < t_{n+1}$. As in the proof of (I), it follows that

$$|A(t, \lambda)| \leqq (1/t\lambda^{\frac{1}{2}}) \int_0^t \omega(s, \pi/\lambda^{\frac{1}{2}})\, d\lambda + O(\lambda^{-3/2}).$$

In view of the definition of $K(\lambda)$, this completes the proof of (9).

9. Proof of (II′) and (II″). Under the conditions of (II′), the arguments leading to (28) hold. Similar arguments show that $2yy'/(\lambda y^2 + y'^2)$, up to a correction term $O(\lambda^{-2})$, is

$$(\nu_k^{\frac{1}{2}}/\lambda)\,(\sin 2\nu_k^{\frac{1}{2}}(s - t_k) - q(s)\lambda^{-1} \cos^2 \nu_k^{\frac{1}{2}}(s - t_k) \sin 2\nu_k^{\frac{1}{2}}(s - t_k)),$$

as $\lambda \to \infty$, uniformly for $t_k \leqq s \leqq t_{k+1}$. Consequently, $A(t, \lambda)$, up to a correction term $O(\lambda^{-2})$, is the sum of the three expressions

$$-\sum_{k=1}^{n} (\nu_k^{\frac{1}{2}}/\lambda t) \int_{t_k}^{t_{k+1}} q(s) \sin 2\nu_k^{\frac{1}{2}}(s - t_k)\, ds, \tag{42}$$

$$\sum_{k=1}^{n} (\nu_k^{\frac{1}{2}}/2\lambda^2 t) \int_{t_k}^{t_{k+1}} q^2(s) \sin 2\nu_k^{\frac{1}{2}}(s - t_k)\, ds,$$

$$\sum_{k=1}^{n} (\nu_k^{\frac{1}{2}}/4\lambda^2 t) \int_{t_k}^{t_{k+1}} q^2(s) \sin 4\nu_k^{\frac{1}{2}}(s - t_k)\, ds.$$

Integrations by parts show that (42) is

$$(1/2\lambda t) \sum_{k=1}^{n} \{q(t_{k+1}) - q(t_k) - \int_{t_k}^{t_{k+1}} q'(s) \cos 2\nu_k^{\frac{1}{2}}(s - t_k)\, ds\},$$

which is

$$O(1/\lambda t) - (1/2\lambda t) \sum_{k=1}^{n} \int_{t_k}^{t_{k+1}} q'(s) \cos 2\nu_k^{\frac{1}{2}}(s - t_k)\, ds, \tag{43}$$

since $q(t)$ is bounded. Integrations by parts also show that both expressions in the formula lines following (42) are $O(1/\lambda^2)$, as $t \to \infty$, for every fixed large λ. Hence,

$$|A(t,\lambda)| \leqq (1/2\lambda t) \int_0^t \omega^*(s, \pi/\lambda^{\frac{1}{2}})\, ds + O(1/\lambda^2),$$

as $t \to \infty$, for large λ. The relation (10) follows.

Under the conditions of (II''), another integration by parts can be applied to each of the integrals in (43), thereby yielding the inequality

$$|A(t,\lambda)| \leqq O(1/\lambda^{3/2} t) \int_0^t \omega^{**}(s, \pi/\lambda^{\frac{1}{2}})\, ds + O(1/\lambda^2),$$

thus completing the proof of (II'').

Appendix.

It was shown in [7], p. 137, that if q satisfies (13) and if $|\lambda| > M$, then a necessary condition for λ to be in the complement of S', is that there exist a number $l > 0$ and a pair of solutions $x_1(t)$, $x_2(t)$ of (1) satisfying (8). It is the purpose of this Appendix to show that this necessary condition is not sufficient. To this end, the example in [13], p. 269, will be refined.

Let the seqeuence of numbers $m_0 < m_1 < \cdots$ be defined as follows:

$m_0 = 0$, $m_{2n+1} = m_{2n} + 2(n+1)\pi$ and $m_{2n+2} = m_{2n+1} + 2(n+1)\pi$ for $n \geqq 0$.

This sequence has the properties that m_n is an even integral multiple of π and that

$$t^{-1} \int_0^t \epsilon(s)\, ds \geqq 1 \text{ if } t > 0, \tag{44}$$

and $\epsilon(s)$ is 2 or 0 according as $m_{2n} \leqq s < m_{2n+1}$ or $m_{2n+1} \leqq s < m_{2n+2}$ for $n = 0, 1, \cdots$.

In terms of $\epsilon(s)$, the sequence $m_0, m_1, \cdots$ and a number $\alpha > 0$, define a function $\chi(s)$, where $0 \leqq s < \infty$, by placing

$$\chi(s) = \alpha\epsilon(s)(1 - \cos s). \tag{45}$$

It is easily seen that $\chi(s)$ possesses a continuous derivative (even at $s = m_0, m_1, \cdots$). It follows from [13], p. 268, that if

$$q(t) = -\epsilon(t)\{2\alpha^2(1-\cos t)^2 \cos^2 t - 3\alpha \sin t + 2\alpha \sin 2t\} \tag{46}$$

and $\lambda = 1$, then (1) has

$$x = x(t) = \exp\Big(\int_0^t \chi(s)\cos s\, ds\Big)\cos t \tag{47}$$

as a solution.

From (45), it is seen that

$$\int_0^t \chi(s)\cos s\, ds = -\alpha \int_0^t \epsilon(s)\,(\tfrac{1}{2} + \tfrac{1}{2}\cos 2s - \cos s)\, ds;$$

and from (44), that

$$\int_0^t \chi(s)\cos s\, ds \leqq -\tfrac{1}{2}\alpha t + O(1)$$

as $t \to \infty$. Hence, $x^2(t) \leqq$ const. $e^{-\alpha t}$ and $x'^2(t) \leqq$ Const. $e^{-\alpha t}$, by (47). Thus, if $x_1(t)$ is defined as a suitable multiple of (47), the first inequality in (8) is satisfied by $l = \alpha > 0$. If $x_2(t)$ is any solution linearly independent of $x_1(t)$, the second inequality of (8) holds with $l = \alpha$; cf., e. g., [7], p. 138.

If $\alpha > 0$ is sufficiently small, it follows from (46) that $|q| \leqq M < 1$. Hence, in order to complete the proof, it remains to show that $\lambda = 1$ is in S' in the case (46). Actually, it will be shown that S' contains the half-line $0 \leqq \lambda < \infty$, by proving that (12) holds for every $\lambda > 0$ and $\Delta > 0$.

The Sturm comparison theorem implies that if $T > m_{2n}$, then $N(T, \lambda + \Delta) - N(T, \lambda)$ is not less than $[N(m_{2n}, \lambda + \Delta) - N(m_{2n-1}, \lambda + \Delta)] - [N(m_{2n}, \lambda) - N(m_{2n-1}, \lambda)] - 2$. On the other hand, since $q(t) = 0$ for $m_{2n-1} \leqq t < m_{2n}$, it follows from (1) that

$$N(m_{2n}, \lambda + \Delta) - N(m_{2n-1}, \lambda + \Delta) \geqq (\lambda + \Delta)^{\frac{1}{2}}(m_{2n} - m_{2n-1})/\pi - 2$$

and

$$N(m_{2n}, \lambda) - N(m_{2n}, \lambda) \leqq \lambda^{\frac{1}{2}}(m_{2n} - m_{2n-1})/\pi + 2.$$

It is seen, from the definition of the sequence $m_0, m_1, \cdots$, that if $T > m_{2n}$, then

$$N(T, \lambda + \Delta) - N(T, \lambda) \geqq ((\lambda + \Delta)^{\frac{1}{2}} - \lambda^{\frac{1}{2}})2n - 6.$$

This completes the proof.

A slight modification of the above example leads to the following: The function

$$x = x(s) = \exp\Big(-\int_0^t \chi(s)\, ds\Big) \tag{48}$$

is a solution of (1) if $\lambda = 0$ and

$$q(t) = -\epsilon(t)(2\alpha^2 \sin t - \alpha \cos t). \tag{49}$$

It is easily seen that (48) satisfies $x^2 + x'^2 = O(e^{-2\alpha t})$, as $t \to \infty$. Hence, there exist linearly independent solutions satisfying (8) with $l = 2\alpha > 0$. Furthermore, in the case (49), the differential equation (1) is oscillatory for $\lambda > 0$, since $q(t) = 0$ for $m_{2n-1} \leqq s < m_{2n-1} + 2n\pi$. Since (48) has no zeros, it follows that $\lambda = 0$ is the least point of S'; [5]. This is in contrast to (47) which belongs to an interior point of S' in the case (46).

THE JOHNS HOPKINS UNIVERSITY.

REFERENCES.

[1] G. Borg, "Eine Umkehrung der Sturm-Liouvillesche Eigenwertaufgabe. Bestimmung der Differentialgleichung durch die Eigenwerte," *Acta Mathematica*, vol. 78 (1946), pp. 319-334.

[2] F. Riesz, "Ueber die Fourierkoeffizienten einer stetigen Funktion von beschränkter Schwankung," *Mathematische Zeitschrift*, vol. 2 (1918), pp. 312-315.

[3] P. Hartman, "On the spectra of slightly disturbed linear oscillators," *American Journal of Mathematics*, vol. 71 (1949), pp. 71-79.

[4] ———, "A characterization of the spectra of one-dimensional wave equations," *ibid.*, vol. 71 (1949), pp. 915-920.

[5] ——— and C. R. Putnam, "The least cluster point of the spectrum of boundary value problems," *ibid.*, vol. 70 (1948), pp. 849-855.

[6] C. R. Putnam, "The cluster spectra of bounded potentials," *ibid.*, vol. 71 (1949), pp. 612-620.

[7] ———, "On isolated eigenfunctions associated with bounded potentials," *ibid.*, vol. 72 (1950), pp. 135-147.

[8] S. Wallach, "The spectra of periodic potentials," *ibid.*, vol. 70 (1948), pp. 842-848.

[9] H. Weyl, "Ueber gewöhnliche Differentialgleichungen mit Singularitäten und die zugehörigen Entwicklungen willkurlicher Funktionen," *Mathematische Annalen*, vol. 68 (1910), pp. 222-269.

[10] ———, "Die Gibbs'sche Erscheinung in der Theorie der Kugelfunktionen," *Rendiconti del Circolo Matematico di Palermo*, vol. 29 (1910), pp. 308-323.

[11] A. Wintner, "The adiabatic linear oscillator," *American Journal of Mathematics*, vol. 68 (1946), pp. 385-397.

[12] ———, "Stability and spectrum in the wave mechanics of lattices," *Physical Review*, vol. 72 (1947), pp. 81-82.

[13] ———, "Asymptotic integration of the adiabatic oscillator," *American Journal of Mathematics*, vol. 69 (1947), pp. 251-272.

[14] ———, "On the smallness of isolated eigenfunctions," *ibid.*, vol. 71 (1949), pp. 603-611.

ON NON-UNIFORMLY QUASI-ANALYTIC FUNCTIONS.*

By I. I. HIRSCHMAN, JR.

1. Introduction. In previous papers, see [1] and [2], various results have been established concerning functions which belong to quasi-analytic classes $-\infty < x < \infty$. For example, it has been shown that if

$$(1) \quad |f^{(n)}(x)| \leqq Ak^n n! [\log(n+e)]^n, \quad -\infty < x < \infty; n = 0, 1, \cdots,$$

and if

$$(2) \quad f(x) = O(\exp[-\exp\exp\{\pi(1+\epsilon)/2\}kx]), \quad x \to +\infty,$$

where $\epsilon > 0$, then $f(x) \equiv 0$. It is natural to inquire whether similar conclusions can be drawn when the inequalities (1) are replaced by the more general inequalities

$$(3) \quad |f^{(n)}(x)| \leqq A[k(x)]^n n! [\log(n+e)]^n, \quad -\infty < x < \infty, n = 0, 1, \cdots.$$

In the present paper we shall obtain some results of this kind. We shall, for example, show that if $f(x)$ satisfies the inequalities (3) where $k'(x)[k(x)]^{-2} = o(1)$ as $x \to +\infty$, and if

$$(4) \quad f(x) = O(\exp[-\exp\exp\tfrac{1}{2}\pi(1+\epsilon)\int_0^x k(t)\,dt]), \quad x \to +\infty$$

with $\epsilon > 0$, then $f(x) \equiv 0$. This result reduces to the previous one when $k(x)$ is constant. The method of demonstration, which is very simple, depends upon conformal mapping of curvilinear strips.

2. Preliminary results. We shall take $z = x + iy$ and $w = u + iv$ throughout.

THEOREM 2a. *Suppose that $f(x)$ is infinitely differentiable in a neighborhood of $x = a$, $\phi(z)$ is analytic for $|z - \alpha| < \rho$, $\phi(\alpha) = a$, and $\phi(x)$ is real for $|x - \alpha| < \rho$. If*

1. $|f^{(n)}(a)| \leqq Ak^n M_n$, $n = 0, 1, \cdots$;
2. $M_n q^n/n!$ *is non decreasing for a fixed* $q > 0$;
3. $|\phi(z) - a| \leqq r$ *for* $|z - \alpha| \leqq \rho$;
4. $l = q\rho^{-1} \operatorname{Max}[1, krq^{-1}]$,

then for $F(x) = f[\phi(x)]$ we have $|F^{(n)}(\alpha)| \leqq nAl^n M_n$, $n = 1, 2, \cdots$.

* Received April 26, 1950

Formally,

$$f(x) = \sum_{j=0}^{\infty} [f^{(j)}(a)/j!](x-a)^j, \quad F(x) = \sum_{j=0}^{\infty} [f^{(j)}(a)/j!][\phi(x)-a]^j,$$

and thus

$$F^{(n)}(\alpha) = \sum_{j=1}^{n} [f^{(j)}(a)/j!](d/dx)^n[\phi(x)-a]^j{}_{x=\alpha}, \qquad n \geqq 1, \tag{1}$$

(the terms corresponding to $j=0$ and $j>n$ are zero). It is easily seen that equation (1) is valid generally. Cauchy's formula yields that

$$(d/dx)^n[\phi(x)-a]^j{}_{x=\alpha} = (2\pi i)^{-1} n! \int_{|z-\alpha|=\rho} \{[\phi(z)-a]^j/[z-\alpha]^{n+1}\}dz,$$

$$|(d/dx)^n[\phi(x)-a]^j{}_{x=\alpha}| \leqq n! r^j \rho^{-n}. \tag{2}$$

Inserting inequality (2) in equation (1) we obtain

$$|F^{(n)}(\alpha)| \leqq \sum_{j=1}^{n} [M_j q^j/j!][An!\rho^{-n}k^j q^{-j} r^j],$$

$\leqq nAl^n M_n$, as desired.

THEOREM 2b. *If*

1. $\theta(x)$ *is defined for* $-\infty < x < \infty$; $\theta(x) > 0$; $\theta'(x) = o(1)$ *as* $x \to +\infty$;
2. S_z *is the strip* $|y| < \theta(x)$, $-\infty\, x < \infty$;
3. S_w *is the strip* $|v| < 1$, $-\infty < u < \infty$;
4. $\psi(z)$ *maps* S_z *conformally onto* S_w, $\psi(0) = 0$, $\psi'(0) > 0$, *and* $\phi(w)$ *is inverse to* $\psi(z)$,

then for $0 < \lambda < 1$ *we have*

A. $\limsup_{x_1 \to +\infty} \operatorname{Max}_{|x-x_1| \leqq \lambda\theta(x_1)} |\psi(z) - \psi(x_1)| = \lambda,$

B. $\limsup_{u_1 \to +\infty} \operatorname{Max}_{|w-u_1| \leqq \lambda} |\phi(w) - \phi(u_1)|/\theta[\phi(u_1)] = \lambda,$

C. $\psi(x) \sim \int_0^x [\theta(t)]^{-1} dt, \qquad x \to +\infty.$

Let $\eta(x_1, x) = \theta[x_1 + \theta(x_1)x]/\theta(x_1)$ and let $S(x_1)$ be the strip $|y| < \eta(x_1, x)$, $-\infty < x < \infty$. It may then be verified that the function $\psi(x_1, z) = \psi[\theta(x_1)z + x_1] - \psi(x_1)$ maps $S(x,)$ onto S_w conformally and that $\psi(x_1, 0) = 0$, $\psi'(x_1, 0) > 0$. It follows from assumption 1 that $\lim_{x_1 \to +\infty} \eta(x_1, x) = 1$ uniformly for x in any finite interval. By a well known theorem on convergent sequences of domains, see [3; § 123], it follows that

$$\lim \psi[\theta(x_1)z + x_1] - \psi(x_1) = z, \qquad (x_1 \to \infty), \tag{3}$$

uniformly for z in any compact subset of the strip $|y| < 1$, $-\infty < x < \infty$. Conclusions A. and B. are immediate consequences of this fact. To demonstrate conclusion C. we choose a sequence $\{x_n\}_0^\infty$, $0 = x_0 < x_1 < x_2 < \cdots$, such that $r' < (x_{n+1} - x_n)/\theta(x_n) < r''$, $n = 0, 1, \cdots$, for some numbers $0 < r' < r'' < \infty$. It follows from equation (3) that

$$\psi(x_{n+1}) - \psi(x_n) \sim (x_{n+1} - x_n)/\theta(x_n).$$

From assumption 1 we may deduce that $(x_{n+1} - x_n)/\theta(x_n) \sim \int_{x_n}^{x_{n+1}} [\theta(t)]^{-1} dt$. Hence summing $\psi(x_n) \sim \int_0^{x_n} [\theta(t)]^{-1} dt$. Since any value of x lies between two values of the sequence $\{x_n\}_0^\infty$ and since $\int_{x_n}^{x_{n+1}} [\theta(t)]^{-1} dt = O(1)$ as $n \to \infty$, our desired result follows.

LEMMA 2c. *We have*

$$\operatorname*{Max}_{|x - x_1| \leqq 1, -\infty < x_1 < \infty} |\, [z^2 + c^2]^{\frac{1}{2}} - [x_1^2 + c^2]^{\frac{1}{2}} \,| = 1 + R(c),$$

where $R(c) = o(1)$ *as* $c \to +\infty$.

The proof of this simple result is left to the reader.

3. Applications. THEOREM 3a. *Let*

1. $\{M_n\}_0^\infty$ *be a logarithmically convex sequence of non-negative constants such that* $\limsup [n!/M_n]^{1/n} < \infty$;
2. $T(r) = \operatorname*{Max}_{n \geqq 0} r^n/M_n$;
3. $\int_1^\infty [\log T(r)] r^{-2} dr = \infty$, $t = (2/\pi) \int_1^{\mathfrak{H}(t)} [\log t(r)] r^{-2} dr$;
4. $k(x) > 0$, $0 \leqq x \leqq \infty$, $k'(x)[k(x)]^{-2} = o(1)$, $x \to +\infty$;
5. $|f^{(n)}(x)| \leqq A M_n [k(x)]^n$, $0 \leqq x < \infty$; $n = 0, 1, \cdots$.

Then if $f(x) = O(\exp\{-\mathfrak{H}[(1+\epsilon)\int_0^x k(t)\,dt]\})$ *as* $x \to +\infty$, *where* $\epsilon > 0$, *we have* $f(x) \equiv 0$.

By logarithmic convexity we have that $M_{n-1} \leqq M_n^{(n-1)/n} M_0^{1/n}$, and hence that $[M_{n-1}/(n-1)!][M_n/n!]^{-1} \leqq n M_0^{1/n} M_n^{-1/n}$. It follows from assumption 1 that $\limsup_{n \to \infty} n M_0^{1/n} M_n^{-1/n} < \infty$, and thus that

$$q = \operatorname*{Max}_{n \geqq 1} [M_{n-1}/(n-1)!][M_n/n!]^{-1}$$

is finite. Clearly, $M_n q^n/n!$ is non-decreasing.

We define $\theta(x) = q/k(0)$ for $-\infty < x \leqq 0$ and $\theta(x) = q/k(x)$ for $0 < x < \infty$. Because of assumption 4. we have $\theta'(x) = o(1)$ as $x \to +\infty$. Let S_z, S_w, and $\psi(z)$ and $\phi(w)$ be defined as in Theorem 2c. Consider the function $g(u) = f[\phi(u)]$. By Theorem 2a and conclusion B of Theorem 2b we see that, given $\delta > 0$, there exists a u_0 such that $|g^{(n)}(u)| \leqq A_0 q^n (1+\delta)^n M_n$ for $u \geqq u_0$, $n = 0, 1, \cdots$. By Theorem 2a and Lemma 2c, given $\delta_1 > \delta$, there exists a value of c such that if $h(t) = g[(t^2 + c^2)^{\frac{1}{2}}]$, then

$$|h^{(n)}(t)| \leqq A_1 q^n (1+\delta_1)^n M_n, \qquad -\infty < t < \infty\,;\ n = 0, 1, \cdots.$$

We have $g[\psi(x)] = f(x)$, and so, by assumption,

$$g[\psi(x)] = O(\exp\{-\mathfrak{H}[(1+\epsilon)\int_0^x k(t)\,dt]\})$$

as $x \to +\infty$. By conclusion C of Theorem 2c we have for $0 < \epsilon_0 < \epsilon$ and x large that

$$(1+\epsilon)\int_0^x k(t)\,dt > (1+\epsilon_0) q\psi(x).$$

It follows that

$$g(x) = O(\exp\{-\mathfrak{H}[(1+\epsilon_0)qx]\})$$

as $x \to +\infty$. If $0 < \epsilon_1 < \epsilon_0$, then

$$h(t) = O(\exp\{-\mathfrak{H}[(1+\epsilon_1)qt]\})$$

as $t \to +\infty$. We may suppose that $(1+\delta_1)/(1+\epsilon_1) < 1$. It then follows from Theorem 4 of [1] that $h(t) \equiv 0$, hence $f(x) \equiv 0$.

THEOREM 3b. *Let*

1. $M_n = n!\,\nu(n)^n$, *where* $\nu'(x) \geqq 0$, *and* $x\nu'(x)/\nu(x) = o(1)$, $x \to \infty$;
2. $T(r) = \operatorname{Max}_{n \geqq 0} r^n/M_n$;
3. $\int_1^\infty \log T(r) r^{-2}\,dr = \infty$, $H(t) = (2/\pi)\int_1^t \log T(r) r^{-2}\,dr$;
4. $k(x) > 0$ $(0 \leqq x < \infty)$, $k'(x)[k(x)]^{-2} = o(1)$, $x \to +\infty$;
5. $|f^{(n)}(x)| \leqq A[k(x)]^n M_n$, $0 \leqq x < \infty$; $n = 0, 1, \cdots$;
6. $Z(u)$ *be the number of zeros of* $f(x)$ *for* $0 \leqq x \leqq u$, *counted according to their multiplicities.*

Then if $f(x) \not\equiv 0$, *we have*

$$\limsup_{x \to +\infty} H[Z(x)] \Big/ \int_0^x k(t)\,dt \leqq 1.$$

The demonstration of this theorem is just like that of Theorem 3a, except that we use Theorem 4b of [2] in place of Theorem 4 of [1].

It is possible to show that the conclusion of Theorem 3a is false for $\epsilon < 0$. One can employ for this purpose the counter-example given in [1], together with Theorem 2a and conclusions A. and C. of Theorem 2b. Similarly, Theorem 3b cannot be greatly improved.

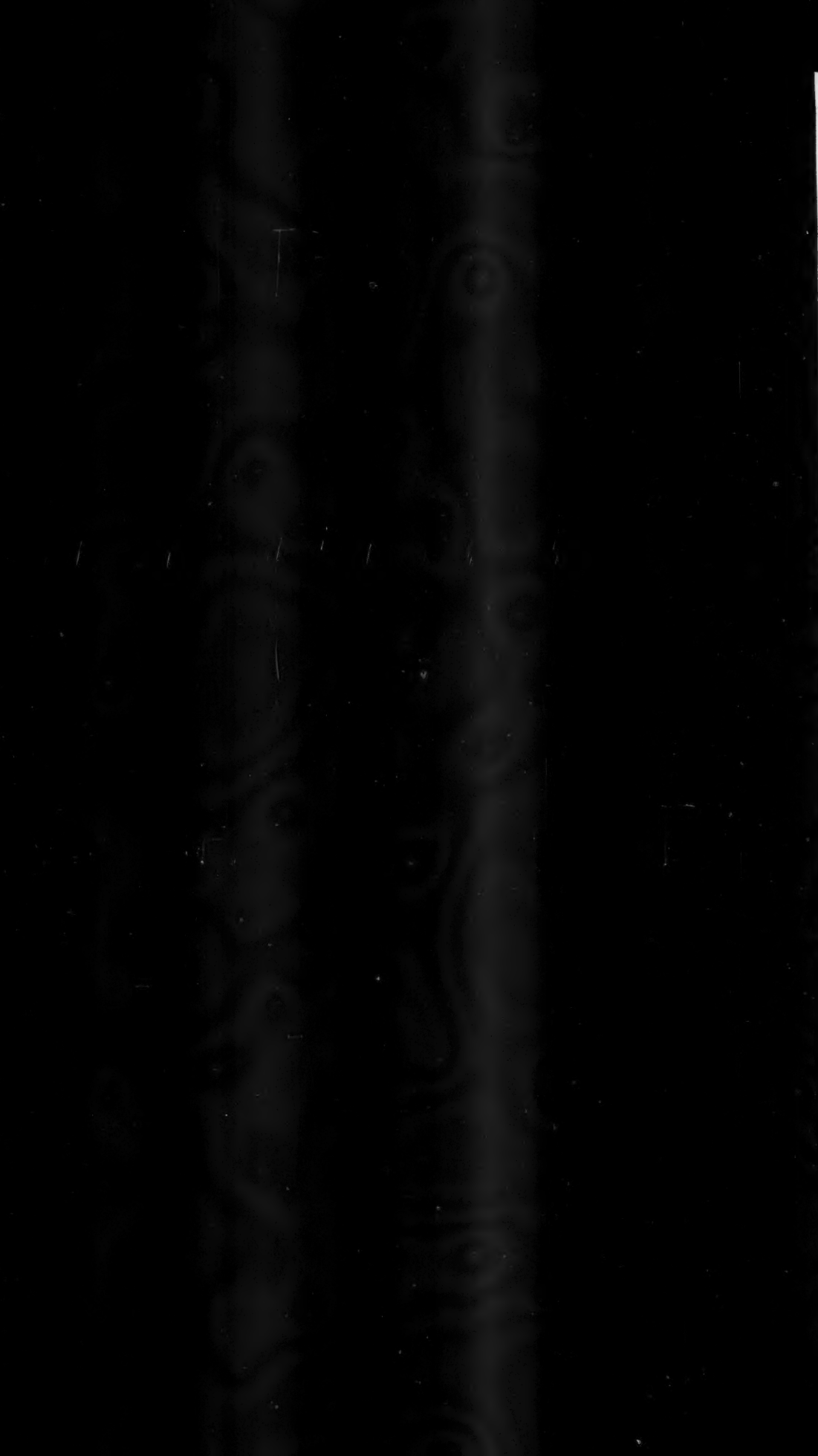
WASHINGTON UNIVERSITY.

BIBLIOGRAPHY.

[1] I. I. Hirschman, Jr., "On the behaviour of Fourier transforms at infinity and on quasi-analytic classes of functions," *American Journal of Mathematics*, vol. 72 (1950), pp. 200-213.

[2] ———, "On the distributions of the zeros of functions belonging to certain quasi-analytic classes," *ibid.*, vol. 72 (1950), pp. 396-406.

[3] C. Carathéodory, *Conformal Representation*, Cambridge, 1941.



American Journal of Mathematics. Edited by R. BRAUER, S. EILENBERG, L. M. GRAVES, D. C. LEWIS, JR. and A. WINTNER. Quarterly. 8vo. Volume LXXII in progress. $7.50 per volume. (Foreign postage, fifty cents.)

American Journal of Philology. Edited by HENRY T. ROWELL, KEMP MALONE, BENJAMIN D. MERITT, JAMES H. OLIVER, EVELYN H. CLIFT, Secretary, and DAVID M. ROBINSON, honorary editor. Quarterly. 8vo. Volume LXXI in progress. $5 per volume. (Foreign postage twenty-five cents.)

Bulletin of the History of Medicine. OWSEI TEMKIN, Editor. Bi-monthly. Volume XXIV in progress. 8vo. Subscription $6 per year. (Foreign postage, fifty cents.)

Bulletin of the Johns Hopkins Hospital. Edited by FREDERICK B. BANG. Monthly. Volume LXXXIV in progress. 8vo. Subscription $6 per year. (Foreign postage, fifty cents.)

ELH. A Journal of English Literary History. Edited by E. T. NORRIS (Managing Editor) and D. C. ALLEN, G. E. BENTLEY, C. P. LYONS, ROY H. PEARCE and E. R. WASSERMAN. Quarterly. 8vo. Volume XVII in progress. $3.50 per volume.

Hesperia. Edited by WILLIAM KURRELMEYER and KEMP MALONE. 8vo. Thirty-four numbers have appeared.

Human Biology: a record of research. CHARLES P. WINSOR, Editor. Quarterly. 8vo. Volume XXII in progress. $5 per volume. (Foreign postage, thirty-five cents.)

Johns Hopkins Studies in Romance Literatures and Languages. H. C. LANCASTER, Editor. 8vo. Seventy-one numbers have been published.

Johns Hopkins University Circular, including the President's Report and Catalogue of the School of Medicine. Eight times yearly. 8vo. $1 per year.

Johns Hopkins University Studies in Archaeology. DAVID M. ROBINSON, Editor. 8vo. Thirty-eight volumes have appeared.

Johns Hopkins University Studies in Education. Formerly edited by FLORENCE E. BAMBERGER. 8vo. Thirty-five numbers have appeared.

Johns Hopkins University Studies in Geology. JOSEPH T. SINGEWALD, JR., Editor. 8vo. Fifteen numbers have been published.

Johns Hopkins University Studies in Historical and Political Science. Under the direction of the Departments of History, Political Economy and Political Science. 8vo. Volume LXVI in progress.

Modern Language Notes. Edited by H. C. LANCASTER, W. KURRELMEYER, K. MALONE, C. R. ANDERSON and D. C. ALLEN. Eight times yearly. 8vo. Volume LXV in progress. $5 per volume. (Foreign postage, fifty cents.)

Reprint of Economic Tracts. Founded by J. H. HOLLANDER. Five series were issued.

The Journal of Geophysical Research. Founded by LOUIS A. BAUER: conducted by M. A. TUVE with the coöperation of eminent investigators. Quarterly. 8vo. Volume LV in progress. $3.50 per volume.

A complete list of publications will be sent on request

Completely Revised And Enlarged

NUMERICAL MATHEMATICAL ANALYSIS

2nd Edition (1950)

By James B. Scarborough

Designed to serve as a textbook and reference work for students and workers in mathematics, statistics, engineering, and the exact natural sciences, Dr. Scarborough's revised and enlarged second edition sets forth logically and clearly the most important principles and processes used for obtaining numerical results. Means for estimating their accuracy are propounded, sufficient theory is given to show their soundness, and the limitations and pitfalls of these methods are pointed out.....$6.00

TABLE OF CONTENTS

CHAPTER

1. The Accuracy of Approximate Calculations
2. Interpolation — *Differences. Newton's Formulas of Interpolation*
3. Interpolation—*Central-Difference Formulas*
4. Interpolation—*LaGrange's Formula. Inverse Interpolation*
5. The Accuracy of Interpolation Formulas
6. Interpolation with two Independent Variables—*Trigonometric Interpolation*
7. Numerical Differentiation and Integration
8. The Accuracy of Quadrature Formulas
9. The Solution of Numerical Algebraic and Transcendental Equations—*Equations in One Unknown. Simultaneous Equations*
10. Graeffe's Root-Squaring Method for Solving Algebraic Equations
11. The Numerical Solution of Ordinary Differential Equations—*Equations of the First Order. Equations of the Second Order and Systems of Simultaneous Equations. The Differential Equations of Exterior Ballistics. Other Methods of Solving Differential Equations Numerically*
12. The Numerical Solution of Partial Differential Equations — *Difference Quotients and Difference Equations. The Method of Iteration. The Method of Relaxation. The Rayleigh-Ritz Method*
13. The Numerical Solution of Integral Equations
14. The Normal Law of Error and the Principle of Least Squares
15. The Precision of Measurements—*Direct Measurements. Indirect Measurements*
16. Empirical Formulas
17. Harmonic Analysis of Empirical Functions

INDEX

ANSWERS TO EXERCISES

The Johns Hopkins Press — Baltimore 18, Maryland